

## SEARCH REQUEST FORM

Requestor's Name: \_\_\_\_\_ Serial Number: \_\_\_\_\_  
Date: \_\_\_\_\_ Phone: \_\_\_\_\_ Art Unit: \_\_\_\_\_

**Search Topic:**

Please write a detailed statement of search topic. Describe specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples or relevant citations, authors keywords, etc., if known. For sequences, please attach a copy of the sequence. You may include a copy of the broadest and/or most relevant claim(s).

## STAFF USE ONLY

Date completed: 06-11-03  
Searcher: Beverly C4994  
Terminal time: 20  
Elapsed time: \_\_\_\_\_  
CPU time: \_\_\_\_\_  
Total time: 23  
Number of Searches: \_\_\_\_\_  
Number of Databases: 1

## Search Site

\_\_\_\_\_ STIC  
\_\_\_\_\_ CM-1  
\_\_\_\_\_ Pre-S

## Type of Search

\_\_\_\_\_ N.A. Sequence  
\_\_\_\_\_ A.A. Sequence  
\_\_\_\_\_ Structure  
\_\_\_\_\_ Bibliographic

## Vendors

\_\_\_\_\_ IG Suite  
\_\_\_\_\_ STN  
\_\_\_\_\_ Dialog  
\_\_\_\_\_ APS  
\_\_\_\_\_ Geninfo  
\_\_\_\_\_ SDC  
\_\_\_\_\_ DARC/Questel  
\_\_\_\_\_ Other CGN

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 9, 2003, 08:36:49 ; Search time 539 Seconds  
(without alignments)  
584.222 Million cell updates/sec

Title: US-09-765-231A-58

Perfect score: 225

Sequence: 1 tgatggttaagtgttcagg.....attagaattttttttttt 225

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 870385 seqs, 699768693 residues

1 number of hits satisfying chosen parameters: 1740770

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA.\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*
- 12: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq.\*
- 13: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*
- 14: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	225	100.0	225	10	US-09-765-231A-58
2	38.2	17.0	640681	10	US-09-790-988-1
3	37.4	16.6	2268	9	US-10-072-130-2
4	36.4	16.2	345	10	US-09-770-791-772
5	36.4	16.2	1479	10	US-09-880-107-1661
6	35.8	15.9	885	9	US-10-198-846-4260
7	35.2	15.6	234	10	US-09-983-965-4454
8	34.6	15.4	1455	10	US-09-917-265-104
9	34.6	15.4	1455	10	US-09-917-265-106
10	34.6	15.4	17029	9	US-10-198-846-10020
11	34.4	15.3	379	9	US-10-198-846-12188
12	34.4	15.3	1132	9	US-09-866-480-98
13	34.4	15.3	2749	9	US-09-813-153-29
14	34.4	15.3	3161	10	US-09-925-301-577
15	34.4	15.3	640681	10	US-09-790-988-1
16	34.2	15.2	824	9	US-09-989-920-13
17	34.2	15.2	30601	9	US-09-373-658-33
18	34.2	15.2	90650	9	US-10-175-523-80
19	34	15.1	584	9	US-10-091-504-2363

C	20	34	15.1	584	10	US-09-764-869-2363	Sequence 2363, Ap
	21	34	15.1	50000	9	US-10-152-724A-20	Sequence 20, Appl
	22	33.8	15.0	412	9	US-09-728-44A-6	Sequence 6, Appl
	23	33.8	15.0	749	10	US-09-925-301-393	Sequence 393, Ap
	24	33.8	15.0	1007	9	US-10-198-846-7198	Sequence 7198, Ap
	25	33.8	15.0	4501	9	US-09-754-853A-10	Sequence 10, Appl
	26	33.8	15.0	4501	9	US-09-754-853A-11	Sequence 11, Appl
	27	33.8	15.0	4508	9	US-09-754-853A-22	Sequence 22, Appl
	28	33.8	15.0	4508	9	US-09-754-853A-23	Sequence 23, Appl
	29	33.8	15.0	4522	9	US-09-754-853A-12	Sequence 12, Appl
	30	33.8	15.0	4522	9	US-09-754-853A-13	Sequence 13, Appl
	31	33.8	15.0	4523	9	US-09-754-853A-8	Sequence 8, Appl
	32	33.8	15.0	4523	9	US-09-754-853A-9	Sequence 9, Appl
	33	33.8	15.0	4523	9	US-09-754-853A-14	Sequence 14, Appl
	34	33.8	15.0	4523	9	US-09-754-853A-15	Sequence 15, Appl
	35	33.8	15.0	4523	9	US-09-754-853A-16	Sequence 16, Appl
	36	33.8	15.0	4523	9	US-09-754-853A-17	Sequence 17, Appl
	37	33.8	15.0	335913	9	US-09-754-853A-2	Sequence 2, Appl
	38	33.8	15.0	335913	9	US-09-754-853A-3	Sequence 3, Appl
	39	33.6	14.9	368	10	US-09-960-352-10867	Sequence 10867, A
	40	33.6	14.9	1406	9	US-10-198-846-5458	Sequence 5458, Ap
	41	33.6	14.9	8009	7	US-08-781-986A-61	Sequence 61, Appl
	42	33.4	14.8	811	9	US-10-078-090-112	Sequence 112, App
	43	33.4	14.8	909	9	US-10-074-095-207	Sequence 207, App
	44	33.4	14.8	909	10	US-09-764-860-207	Sequence 207, App
	45	33.4	14.8	3410	9	US-10-196-063-1	Sequence 1, Appl

#### ALIGNMENTS

##### RESULT 1

US-09-765-231A-58  
; Sequence 58, Application US/09765231A  
; Patent No. US20020119452A1  
; GENERAL INFORMATION:  
; APPLICANT: Searte/Monsanto  
; APPLICANT: Phippard, Deborah  
; APPLICANT: Vasanthakumar, Geetha  
; APPLICANT: Dotson, Stanton  
; APPLICANT: Ma, Xiao-Jun  
; TITLE OF INVENTION: Osteoarthritis tissue-derived nucleic acids, polypeptides,  
; FILE REFERENCE: SO-3221 PR  
; CURRENT APPLICATION NUMBER: US/09/765,231A  
; CURRENT FILING DATE: 2001-01-18  
; NUMBER OF SEQ ID NOS: 82  
; SEQ ID NO 58  
; LENGTH: 225  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-765-231A-58

Query Match	100.0%	Score	225	DB	10	Length	225
Best Local Similarity	100.0%	Pred. No.	1.5e-50				
Matches	225	Conservative	0	Mismatches	0	Indels	0
Gaps	0						
Qy	1	TGATGGTAAAGTTGTTTCAGGCATAAAATTTGAAATAAAATTTAGAGGCTCCATCATATGCT	60				
Db	1	TGATGGTAAAGTTGTTTCAGGCATAAAATTTGAAATAAAATTTAGAGGCTCCATCATATGCT	60				
Qy	61	ATATTGGTTTTACCTTTCAGAAGAATATTTAGTTTCACTCAGGTTTTTCAAGAGCTACGCTG	120				
Db	61	ATATTGGTTTTACCTTTCAGAAGAATATTTAGTTTCACTCAGGTTTTTCAAGAGCTACGCTG	120				
Qy	121	TCCCCCAAAAAACGAAACAAAAACAAACCTTTTTTAAGAGTTGATGGTACTCAT	180				
Db	121	TCCCCCAAAAAACGAAACAAAAACAAACCTTTTTTAAGAGTTGATGGTACTCAT	180				
Qy	181	TTGATCTGCTCTCTGCTGTAATCAATTAGGAATTTTTTTTTTTT	225				
Db	181	TTGATCTGCTCTCTGCTGTAATCAATTAGGAATTTTTTTTTTTT	225				

RESULT 2  
US-09-790-988-1/c  
; Sequence 1, Application US/09790988  
; Patent No. US20020127687A1  
; GENERAL INFORMATION:  
; APPLICANT: SHIGENOBU, SHUJI  
; APPLICANT: WATANABE, HIDEKI  
; APPLICANT: HATTORI, MASAHIRA  
; APPLICANT: SAKAKI, YOSHIYUKI  
; TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS  
; FILE REFERENCE: 081356/0159  
; CURRENT APPLICATION NUMBER: US/09/790,988  
; CURRENT FILING DATE: 2001-02-23  
; PRIOR APPLICATION NUMBER: JP2000-107160  
; PRIOR FILING DATE: 2000-04-07  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: Patent in Ver. 2.1  
; SEQ ID NO 1  
; TYPE: DNA  
; ORGANISM: Buchnera sp.  
US-09-790-988-1

Query Match 17.0%; Score 38.2; DB 10; Length 640681;  
Best Local Similarity 55.7%; Pred. No. 19;  
Matches 73; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 32 AATAAATTATGAGGCTCCATGATATGCTATATGTTTACCTTCAGAGAATATTAG 91  
DB 198302 AATCAAAATTAAATAAAGCTAGTATATTTTCAATTATGACTTTTCAATAATTTTATTTATTAAG 198243

QY 92 TTTTCACTCAGGTTTTTCAAGCTAGCTGTCCTCCCAAAACGAAACAAACAAAAAAC 151  
DB 198242 TAATAATTGTTATTATATAATAACTTTTCACTAAATAAACAATACAAAAA 198183

QY 152 AACCTTTTAA 162  
DB 198182 AAGATTTTAA 198172

RESULT 3  
US-10-072-130-2/c  
; Sequence 2, Application US/10072130  
; Patent No. US20020173022A1  
; GENERAL INFORMATION:  
; APPLICANT: Bandman, Olga  
; Goli, Surya K.  
; Lal, Preeti  
; Corley, Neil C.  
; Zhang, Hong  
; TITLE OF INVENTION: NEW PROTEIN PHOSPHATASE  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/072,130  
; FILING DATE: 05-Feb-2002  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/873,093  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0319 US  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2268 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: 13177  
CLONE: THPIPLB01  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-10-072-130-2

Query Match 16.6%; Score 37.4; DB 9; Length 2268;  
Best Local Similarity 53.0%; Pred. No. 2.9;  
Matches 80; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 29 TTGAATAAATTATGAGGCTCCATGATATGCTATATGTTTACCTTCAGAGAATATT 88  
DB 2195 TTGACATACATTTTGATGTGTAGTGTAGCTATACAGAAATATACAAACCAAGCTGTAT 2136

QY 89 TAGTTTCACTCAGGTTTTTCAAGCTAGCTGTCCTCCCAAAACGAAACAAACAAAA 148  
DB 2135 GAATAATACATAGGTTTTTCAAGCTATGTTTTCATAAGAATACAGAGAAAGGAAAT 2076

QY 149 AACAACTTTTAAAGTTGATGGCTACTCA 179  
DB 2075 AGCCACACATCCATATCTCACAACCTTCAA 2045

RESULT 4  
US-09-770-791-772/c  
; Sequence 772, Application US/09770791  
; Patent No. US20020062014A1  
; GENERAL INFORMATION:  
; APPLICANT: Goriach, Jorn  
; APPLICANT: An, Yong-Qiang  
; APPLICANT: Hamilton, Carol M.  
; APPLICANT: Price, Jennifer L.  
; APPLICANT: Raines, Tracy M.  
; APPLICANT: Yu, Yang  
; APPLICANT: Rameaka, Joshua G.  
; APPLICANT: Page, Amy  
; APPLICANT: Matthew, Abraham V.  
; APPLICANT: Ledford, Brooke L.  
; APPLICANT: Woessner, Jeffrey P.  
; APPLICANT: Haas, William David  
; APPLICANT: Garcia, Carlos A.  
; APPLICANT: Kricker, Maja  
; APPLICANT: Slader, Ted  
; APPLICANT: Davis, Keith R.  
; APPLICANT: Allen, Keith  
; APPLICANT: Hoffman, Neil  
; APPLICANT: Hurban, Patrick  
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis  
; TITLE OF INVENTION: thaliana  
; FILE REFERENCE: 2029 (PARA-018PRV)  
; CURRENT APPLICATION NUMBER: US/09/770,791  
; CURRENT FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: 60/178,480  
; PRIOR FILING DATE: 2000-01-27  
; NUMBER OF SEQ ID NOS: 999  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 772  
; LENGTH: 345  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-770-791-772

```

Query Match      16.2%; Score 36.4; DB 10; Length 345;
Best Local Similarity 53.5%; Pred. No. 2.5;
Matches 76; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY      12  TGTTCAGGCATAAAATTTGAAATAAATATATGAGGCTCCATGATATGCTATATTGGTTTT 71
DB      155  TGTATTACCTAAGCAATAATATGATTTTGAATAAACTAATTTGAATTACTGCATAT 96

QY      72  ACCTTCAGAAAGATATATTAGTTTTCACCTCAGGTTTTTCAAAGCTACGCTGCCCCCAAAA 131
DB      95  GAAATTATAAATCTTTTGTGTTTAACTCTCAAGATGTAATAATTTATGATCTTCACATAAA 36

QY      132  ACGAAACAAACAAAAACAA 153
DB      35  AAAAAAAAAAAAAAAAAAAAA 14

RESULT 5
US-09-880-107-1661
Sequence 1661, Application US/09880107
Patent No. US20020142981A1
GENERAL INFORMATION:
; APPLICANT: Vockley, Darci T.
; APPLICANT: Schert, Uwe
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 1661
; LENGTH: 1479
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 D45288
US-09-880-107-1661

Query Match      16.2%; Score 36.4; DB 10; Length 1479;
Best Local Similarity 61.7%; Pred. No. 4.5;
Matches 58; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

DB      128  AAAAAAGAAACAAACAAAAACAAACCTTTTAAAGAGATTGATGGCTACTCATTTGATCT 187
DB      834  AGAAAGAAACAAACAAACAAACCTTTTAAATTTACTTTGCAACTCAACAGATCT 893

QY      188  GCCTCTCTCTGTAATCAATTAGGAATTTTTT 221
DB      894  CCCTGCGTACTGCTTTTCCAGGAACCTTTACTT 927

RESULT 6
US-10-198-846-4260/c
Sequence 4260, Application US/10198846
Publication No. US2003009974A1
GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220

```

Db 159 TGATGCTTTAGACAGATATAGGTATGAAAAAATTTCTGATAACATGATATATAAAA 100

Qy 64 TTGGTTTACCTTCAGAGAAATATTAGTTTCACTCAGGTTTTT 107

Db 99 ATGTTATTACATGGAAGAATAGTACCCCTCACACAGAGATT 56

RESULT 8

US-09-917-265-104

; Sequence 104, Application US/09917265

; Patent No. US20020052030A1

; GENERAL INFORMATION:

; APPLICANT: Wonderling, Ramani S.

; APPLICANT: Boroughs, Karen L.

; TITLE OF INVENTION: CANINE AND FELINE PROTEINS, NUCLEIC ACID MOLECULES AND USES THERE

; FILE REFERENCE: IM-5

; CURRENT APPLICATION NUMBER: US/09/917,265

; CURRENT FILING DATE: 2001-07-27

; PRIOR APPLICATION NUMBER: 60/223,016

; PRIOR FILING DATE: 2000-08-04

; NUMBER OF SEQ ID NOS: 109

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 104

; LENGTH: 1455

; TYPE: DNA

; ORGANISM: Canis familiaris

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (232)..(897)

; OTHER INFORMATION:

US-09-917-265-104

Query Match 15.4%; Score 34.6; DB 10; Length 1455;

Best Local Similarity 51.8%; Pred. No. 13;

Matches 79; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

Qy 1 TGATGCTTTAGTCTTTTCAGGCATATAAATTTGAAATAAAATTTATGAGGCTCCATGATGCT 60

Db 1298 TTATATAATTTTGTGTTCTCATGAAGCAGTGAATTAATTTATTTATTTATGTTATTT 1357

Qy 61 ATATTGGTTTACCTTCAGAGAAATATTAGTTTCACTCAGGTTTTTCAAAGCTAGCGTG 120

Db 1358 TATTAAAGTATTTATTATCATCAAGTGGATTTGGGATATCTTATGTTCTAAATAAATAATGAT 1417

Qy 121 TCCCCCAAAACGAAACAAACAAACAAACAA 153

Db 1418 TGAGTAGAAAAAATAAATAAATAAATAAATAA 1450

ULT 9

US-09-917-265-106/c

; Sequence 106, Application US/09917265

; Patent No. US20020052030A1

; GENERAL INFORMATION:

; APPLICANT: Wonderling, Ramani S.

; APPLICANT: Boroughs, Karen L.

; TITLE OF INVENTION: CANINE AND FELINE PROTEINS, NUCLEIC ACID MOLECULES AND USES THERE

; FILE REFERENCE: IM-5

; CURRENT APPLICATION NUMBER: US/09/917,265

; CURRENT FILING DATE: 2001-07-27

; PRIOR APPLICATION NUMBER: 60/223,016

; PRIOR FILING DATE: 2000-08-04

; NUMBER OF SEQ ID NOS: 109

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 106

; LENGTH: 1455

; TYPE: DNA

; ORGANISM: Canis familiaris

US-09-917-265-106

Query Match 15.4%; Score 34.6; DB 10; Length 1455;

Best Local Similarity 51.8%; Pred. No. 13;

Matches 79; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

Qy 1 TGATGCTTTAGTCTTTTCAGGCATATAAATTTGAAATAAAATTTATGAGGCTCCATGATGCT 60

Db 158 TTATATAATTTTGTGTTCTCATGAAGCAGTGAATTAATTTATTTATTTATTTATTTATTT 99

Qy 61 ATATTGGTTTACCTTCAGAGAAATATTAGTTTCACTCAGGTTTTTCAAAGCTAGCGTG 120

Db 98 TATTAAAGTATTTATTATCATCAAGTGGATTTGGGATATCTTATGTTCTAAATAAATAATGAT 39

Qy 121 TCCCCCAAAACGAAACAAACAAACAAACAA 153

Db 38 TGAGTAGAAAAAATAAATAAATAAATAAATAA 6

RESULT 10

US-10-198-846-10020/c

; Sequence 10020, Application US/10198846

; Publication No. US2003009974A1

; GENERAL INFORMATION:

; APPLICANT: Lillie, James

; APPLICANT: Xu, Yongyao

; APPLICANT: Wang, Youzhen

; APPLICANT: Steinmann, Kathleen

; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS

; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND

; FILE REFERENCE: MRI-049

; CURRENT APPLICATION NUMBER: US/10/198,846

; CURRENT FILING DATE: 2002-07-18

; PRIOR APPLICATION NUMBER: 60/306,220

; PRIOR FILING DATE: 2001-07-18

; NUMBER OF SEQ ID NOS: 14084

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 10020

; LENGTH: 1747

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: 1, 1747

; OTHER INFORMATION: n = A,T,C or G

US-10-198-846-10020

Query Match 15.4%; Score 34.6; DB 9; Length 1747;

Best Local Similarity 55.4%; Pred. No. 15;

Matches 67; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

Qy 30 TGAATAAATTTATGAGGCTCCATGATATGCTATATTTGTTTACCTTCAGAGAAATATT 89

Db 1390 TAAATAATATATATCAGGCCCATCTCTTTTTTTTTTTTTTTTTTTTTTTTAAAGATGGAGTCTC 1331

Qy 90 AGTTTCACTCAGGTTTTTCAAAGCTAGCGTCCCCCAAAACGAAACAAACAAACAA 149

Db 1330 ACTCTGTCAACCAGGCTGGAGTCTGTATCGAAGAAAAAATAAATAAATAAAGGAA 1271

Qy 150 A 150

Db 1270 A 1270

RESULT 11

US-10-198-846-12188/c

; Sequence 12188, Application US/10198846

; Publication No. US2003009974A1

; GENERAL INFORMATION:

; APPLICANT: Lillie, James

; APPLICANT: Xu, Yongyao

; APPLICANT: Wang, Youzhen

; APPLICANT: Steinmann, Kathleen

; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS

; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND

; FILE REFERENCE: MRI-049

CURRENT APPLICATION NUMBER: US/10/198,846  
CURRENT FILING DATE: 2002-07-18  
PRIOR APPLICATION NUMBER: 60/306,220  
PRIOR FILING DATE: 2001-07-18  
NUMBER OF SEQ ID NOS: 14084  
SOFTWARE: FASTSEQ for Windows Version 4.0  
SEQ ID NO 12188  
LENGTH: 379  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: 378..379  
OTHER INFORMATION: n = A,T,C or G  
US-10-198-846-12188

Query Match 15.3%; Score 34.4; DB 9; Length 379;  
Best Local Similarity 54.8%; Pred. No. 8.7;  
Matches 68; Conservative 0; Mismatches 56; Indels 0; Gaps 0;  
27 ATTTGAAATAAATTATGAGGCTCCATGATATGCTATATTGGTTTACCTTCAGAAGAATA 86  
126 AATTAGCTATATTATCTCTCAATCAAAATCTACTTCAGAGGTAAGTAAAAAAGAAGT 67  
87 TTTAGTTTCACTCAGGTTTTTCAAAGCTAGCTGTCCCCCAAAAAACGAAACAAACAAA 146  
66 TATAGCAATACATAAAATGGAACAAAAGGTATCCCAAAAAAATAAAAAAATAAAAA 7  
147 AAAA 150  
6 AAAA 3

RESULT 12  
US-09-986-480-98  
Sequence 98, Application US/09986480  
Publication No. US20030027999A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
FILE REFERENCE: PS500P1  
CURRENT APPLICATION NUMBER: US/09/986,480  
CURRENT FILING DATE: 2001-11-08  
PRIOR APPLICATION NUMBER: PCT/US00/12788  
PRIOR FILING DATE: 2000-05-11  
PRIOR APPLICATION NUMBER: US 60/134,068  
PRIOR FILING DATE: 1999-05-13  
NUMBER OF SEQ ID NOS: 456  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 98  
LENGTH: 1132  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (153)  
OTHER INFORMATION: n equals a,t,g, or c  
NAME/KEY: SITE  
LOCATION: (260)  
OTHER INFORMATION: n equals a,t,g, or c  
US-09-986-480-98

Query Match 15.3%; Score 34.4; DB 9; Length 1132;  
Best Local Similarity 54.8%; Pred. No. 14;  
Matches 68; Conservative 0; Mismatches 56; Indels 0; Gaps 0;  
30 TGAATAAATAATGAGGCTCCATGATATGCTATATTGGTTTACCTTCAGAAGAATAATT 89  
1009 TGAAGAATAATTGTTATATACAGAAGAGCCAGATGAATATTGTTTAAAGTAGAGCATTT 1068  
90 ACTTTCACTCAGGTTTTTCAAAGCTAGCTGTCCCAAAAAACGAAACAAACAAAAA 149  
1069 CCTTTTATTAAATAAATAAATAAATGTTTACATCCAAAAAATAAAAAAATAAAC 1128

QY 150 ACAA 153  
DB 1129 TCGA 1132

RESULT 13  
US-09-813-153-29  
Sequence 29, Application US/09813153  
Publication No. US20030045459A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: 67 Human secreted proteins  
FILE REFERENCE: P2023  
CURRENT APPLICATION NUMBER: US/09/813,153  
CURRENT FILING DATE: 2001-03-21  
PRIOR APPLICATION NUMBER: US/09/363,044  
PRIOR FILING DATE: 1999-07-29  
PRIOR APPLICATION NUMBER: 60/073,160  
PRIOR FILING DATE: 1998-01-30  
PRIOR APPLICATION NUMBER: 60/073,159  
PRIOR FILING DATE: 1998-01-30  
PRIOR APPLICATION NUMBER: 60/073,165  
PRIOR FILING DATE: 1998-01-30  
PRIOR APPLICATION NUMBER: 60/073,164  
PRIOR FILING DATE: 1998-01-30  
PRIOR APPLICATION NUMBER: 60/073,167  
PRIOR FILING DATE: 1998-01-30  
PRIOR APPLICATION NUMBER: 60/073,162  
PRIOR FILING DATE: 1998-01-30  
PRIOR APPLICATION NUMBER: 60/073,161  
PRIOR FILING DATE: 1998-01-30  
PRIOR APPLICATION NUMBER: 60/073,170  
PRIOR FILING DATE: 1998-01-30  
NUMBER OF SEQ ID NOS: 298  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 29  
LENGTH: 2749  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-813-153-29

Query Match 15.3%; Score 34.4; DB 9; Length 2749;  
Best Local Similarity 51.3%; Pred. No. 20;  
Matches 77; Conservative 1; Mismatches 72; Indels 0; Gaps 0;  
4 TGGTAAGTTGTTTCAGGCAATAAAATTTGAAATAAATATGAGGCTCCATGATATGCTATA 63  
2578 TGATATGATATTGTTGTTGAAAGTTTTTTGTAAAAAATAATTTTACAATGTTTATTTGAAT 2637  
64 TTGGTTTTACCTTCAGAGAATAATTAGTTTTCACCTCAGGTTTTTCAAAGCTACGCTGTCC 123  
2638 GATTTTTTAAATGCTGTGTAATCTATATTGTTGTTTATATATAAATAATTCATTGGCCA 2697  
124 CCCAAAAACGAAACAAACAAAAAACA 153  
2698 AAAAAAAAAAAAAAAAAAAAAAAAAA 2727

RESULT 14  
US-09-925-301-577  
Sequence 577, Application US/09925301  
Patent No. US20020052308A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
FILE REFERENCE: PA106  
CURRENT APPLICATION NUMBER: US/09/925,301  
CURRENT FILING DATE: 2001-08-10  
PRIOR APPLICATION NUMBER: PCT/US00/05882  
PRIOR FILING DATE: 2000-03-08  
PRIOR APPLICATION NUMBER: 60/124,270  
PRIOR FILING DATE: 1999-03-12





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 9, 2003, 08:36:50 ; Search time 1054 Seconds  
(without alignments)  
3457.291 Million cell updates/sec

Title: US-09-765-231a-58

Perfect score: 225

Sequence: 1 tcatggttaagtgtttcagg.....attaggaattttttttttttt 225

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

ll number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pin:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	196	87.1	341	9 AA502552	AA502552 ng62e06.s
2	117	52.0	324	9 A1247782	A1247782 qb61a07.x
3	117	52.0	362	9 AA682512	AA682512 z119a01.s
4	48.8	21.7	928	17 CNS00DKY	AL071885 Drosophila
5	43	19.1	773	17 AQ781761	AQ781761 HS_3122_A
6	41.8	18.6	964	17 CNS058MA	AL326107 Tetraodon

C	7	41.2	18.3	1101	17	CNS00D90
	8	41	18.2	290	10	AW504318
	9	41	18.2	1562	11	BC022863
	10	40.6	18.0	595	17	AZ523166
	11	40.6	18.0	957	17	CNS015W7
	12	40.4	18.0	525	10	AW381028
	13	40	17.8	475	10	BE325739
	14	39.8	17.7	208	9	AU071524
	15	39.8	17.6	647	17	B83740
	16	39.6	17.6	360	9	AL750985
	17	39.6	17.6	409	14	B0451684
	18	39.6	17.6	524	17	AZ055107
	19	39.6	17.6	941	17	BH133309
	20	39.4	17.5	189	13	BH351968
	21	39.2	17.4	461	13	B1815488
	22	39	17.3	521	10	AW952962
	23	39	17.3	854	12	BG209526
	24	39	17.3	905	17	BH156645
	25	38.8	17.2	500	9	AU088212
	26	38.8	17.2	925	17	CNS01310
	27	38.6	17.2	400	12	BF014395
	28	38.6	17.2	405	12	BF014378
	29	38.6	17.2	419	12	BF014182
	30	38.6	17.2	503	10	BE161482
	31	38.4	17.1	632	13	BM160134
	32	38.4	17.1	716	13	B1294873
	33	38.4	17.1	836	17	CNS02M02
	34	38.4	17.1	1101	17	CNS0039G
	35	38.4	17.1	1101	17	CNS0039W
	36	38.2	17.0	450	17	AZ294776
	37	38.2	17.0	537	17	AZ294781
	38	38.2	17.0	543	17	AZ113230
	39	38.2	17.0	983	17	CNS00KNG
	40	38	16.9	315	14	C92155
	41	38	16.9	442	9	AA134993
	42	38	16.9	671	17	AZ111894
	43	38	16.9	765	17	AQ422165
	44	38	16.9	952	17	CNS016RS
	45	37.8	16.8	253	14	BM882716

#### ALIGNMENTS

RESULT 1  
AA502552  
LOCUS ng62e06.s1 NCI\_CGAP\_Lip2 Homo sapiens cDNA clone IMAGE:939394, mRNA  
DEFINITION sequence.  
ACCESSION AA502552  
VERSION AA502552.1 GI:2237519  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 341)  
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R. Emmert-Buck, M.D., Ph.D.  
CDNA Library Preparation: David B. Krizman, Ph.D.  
DNA Sequencing Arrayed by: Greg Lennon, Ph.D.  
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: [www.bio.llnl.gov/bbrp/image/image.html](http://www.bio.llnl.gov/bbrp/image/image.html)  
Insert Length: 1183 Std Error: 0.00  
Seq primer: -40ml3 fwd. ET from Amersham

High quality sequence stop: 310.  
Location/Qualifiers  
1. .341  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:939394"  
/clone\_lib="NCI\_CGAP\_Lip2"  
/tissue\_type="liposarcoma"  
/lab\_host="DH108"  
/note="Vector: pAMP10; mRNA made from liposarcoma, cDNA made by oligo-dr priming. Non-directionally cloned. Size-selected on agarose gel, average insert size 600 bp. Reference: Krizman et al. (1996) Cancer Research 56:5380-5383."

BASE COUNT 112 a 59 c 59 g 111 t

ORIGIN

Query Match 87.1%; Score 196; DB 9; Length 341;  
Best Local Similarity 99.5%; Pred. No. 9.6e-30;  
Matches 207; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

19 GGCATAAAATTCGAATAAATATGAGGCTCCATGATATGCTATATGTTTACCTTCA 78  
|||||  
20 GGCATAAAATTCGAATAAATATGAGGCTCCATGATATGCTATATGTTTACCTTCA 67  
|||||

79 GAAGAATATTTAGTTTCACTCAGCTTTTCCAAAGCTACGCTGCCCAAAACGAAAC 138  
|||||

68 GAAGATATTTAGTTTCACTCAGCTTTTCCAAAGCTACGCTGCCCAAAACGAAAC 127  
|||||

139 AAAAC-AAAAAACAACCTTTTAAAGAGTTGATGGCTACTCATTTGATCTGCTCTCTG 197  
|||||

128 AAAACAAAAAACAACCTTTTAAAGAGTTGATGGCTACTCATTTGATCTGCTCTCTG 187  
|||||

198 CTGAATCAATTAGGAATTTTTTTTTT 225  
|||||

188 CTGAATCAATTAGGAATTTTTTTTTT 215  
|||||

RESULT 2  
AI247782 324 bp mRNA linear EST 01-DEC-1998  
LOCUS qh61a07.x1 Soares fetal liver spleen INFLS\_S1 Homo sapiens cDNA  
DEFINITION clone IMAGE:1849140 3', mRNA sequence.

ACCESSION AI247782  
VERSION AI247782.1 GI:3843179  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 324)  
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgaps@email.nih.gov  
This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.  
Insert Length: 821 Std Error: 0.00  
Seq primer: -40UP from Gibco  
High quality sequence stop: 317.  
Location/Qualifiers  
1. .324  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:1849140"  
/clone\_lib="Soares\_fetal\_liver\_spleen\_INFLS\_S1"  
/sex="male"  
/dev\_stage="20 week-post conception fetus"  
/lab\_host="DH108 (ampicillin resistant)"  
/note="Organ: Liver and Spleen; Vector: p7T73D (Pharmacia) with a modified polylinker; Site\_1: Pac I; Site\_2: Eco RI; This is a subtracted version of the original Soares fetal liver spleen INFLS library. 1st strand cDNA was primed with a Pac I - oligo(dT) primer [5', AACTGGAAGATTAATTAAGATCTTTTTTTTTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac I and cloned into the Pac I and Eco RI sites of the modified p7T73 vector. Library went through one round of normalization. Library

This is a subtracted version of the original Soares fetal liver spleen INFLS library. 1st strand cDNA was primed with a Pac I - oligo(dT) primer [5', AACTGGAAGATTAATTAAGATCTTTTTTTTTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac I and cloned into the Pac I and Eco RI sites of the modified p7T73 vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo."

BASE COUNT 105 a 64 c 50 g 105 t

ORIGIN

Query Match 52.0%; Score 117; DB 9; Length 324;  
Best Local Similarity 100.0%; Pred. No. 5.7e-14;  
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

109 AAAGCTACGCTGCTCCCAAAACGAAACAAACAAACAAACCAACCTTTTAAAGATTG 168  
|||||

1 AAAGCTACGCTGCTCCCAAAACGAAACAAACAAACCAACCTTTTAAAGATTG 60  
|||||

169 ATGGCTACTCAATTTGATCTGCTCTCTGCTCAATCAATTAGGAATTTTTTTTTT 225  
|||||

61 ATGGCTACTCAATTTGATCTGCTCTCTGCTCAATCAATTAGGAATTTTTTTTTT 117  
|||||

RESULT 3  
AA682512 362 bp mRNA linear EST 19-DEC-1997  
LOCUS zil9a01.s1 Soares fetal liver spleen INFLS\_S1 Homo sapiens cDNA  
DEFINITION clone IMAGE:431208 3', mRNA sequence.

ACCESSION AA682512  
VERSION AA682512.1 GI:2669793  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 362)  
AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S., Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Schellenberg, K., Steptoe, W., Tan, F., Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.  
WashU-NCI human EST Project  
Unpublished (1997)  
Contact: Wilton RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: estowatson.wustl.edu  
This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.  
Seq primer: -40m13 fwd. ET from Amersham  
High quality sequence stop: 308.  
Location/Qualifiers  
1. .362  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:431208"  
/clone\_lib="Soares\_fetal\_liver\_spleen\_INFLS\_S1"  
/sex="male"  
/dev\_stage="20 week-post conception fetus"  
/lab\_host="DH108 (ampicillin resistant)"  
/note="Organ: Liver and Spleen; Vector: p7T73D (Pharmacia) with a modified polylinker; Site\_1: Pac I; Site\_2: Eco RI; This is a subtracted version of the original Soares fetal liver spleen INFLS library. 1st strand cDNA was primed with a Pac I - oligo(dT) primer [5', AACTGGAAGATTAATTAAGATCTTTTTTTTTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac I and cloned into the Pac I and Eco RI sites of the modified p7T73 vector. Library went through one round of normalization. Library

```

BASE COUNT      114 a 71 c 67 g 110 t
ORIGIN
Query Match      52.0%; Score 117; DB 9; Length 362;
Best Local Similarity 100.0%; Pred. No. 5.4e-14;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 109 AAGCTACGCTGTCCTCCCAAAACGAACAAACAAACAAACAAACCTTTTAAAGAGTTG 168
DB 1 AAGCTACGCTGTCCTCCCAAAACGAACAAACAAACAAACAAACCTTTTAAAGAGTTG 60

QY 169 ATGGCTACTCATTTGATCTGCTCTCTCTCTGATCAATAGGAATTTTATTTT 225
DB 61 ATGGCTACTCATTTGATCTGCTCTCTCTCTGATCAATAGGAATTTTATTTT 117

RESULT 4
CNS000DKY
LOCUS
DEFINITION
Drosophila melanogaster genome survey sequence T7 end of BAC #
BACR27A24 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION      AL071865.1 GI:4948170
VERSION
KEYWORDS
SOURCE
ORGANISM
Drosophila melanogaster.
Drosophila melanogaster.
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 928)
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mamooser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila\_bac.htm.
FEATURES
source
1..928
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="BACR27A24"
/clone_lib="RPCI-98"
/note="end : T7"

BASE COUNT      262 a 70 c 84 g 321 t 191 others
ORIGIN
Query Match      21.7%; Score 48.8; DB 17; Length 928;
Best Local Similarity 31.4%; Pred. No. 1.4;
Matches 64; Conservative 59; Mismatches 81; Indels 0; Gaps 0;

QY 22 ATAAATTTTGAATAAATTATGAGGCTCCATGATGCTATATGTTTACCTTCAGAA 81
DB 678 AAAAAATTTTAAWAAWTAATAAAWAAAAATTTAAWTTTTTTTTTTTTTTTATWTATWATAWAAA 737

QY 82 GATATATTAGTTTCACCTCAGGTTTTTCAAGAGCTACGCTGCCCAACAAACAAACAA 141
DB 738 TATWTTWTTWTDGKNNNNAWAAAAWAAAAWAAAAWAAAAWAAAAWAAAAWAAAA 797

```

```

QY 142 ACAAAAAACAACCTTTTAAAGATGTAGGTACTACTATTGATCGCTCTCTCTGTA 201
DB 798 AAAAAWAAAAAAWDDDDDDWAAKAKKKKKKKKKKKKKKKKKKKTKTKTKTKGA 857

QY 202 ATCAATTAGGAATTTTATTTT 225
DB 858 RWWTTTTTTTTTTTTTTTTTTTTTT 881

RESULT 5
A0781761
LOCUS
DEFINITION
HS-3122 A2_D08 T7C CIT Approved Human Genomic Sperm Library D Homo
sapiens Genomic clone Plates=3122 Col=16 Row=G, DNA sequence.
ACCESSION      A0781761
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 773)
REFERENCE
AUTHORS
Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
Hood,L.
TITLE
Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
JOURNAL
MEDLINE
COMMENT
Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
99380589
Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones may be purchased from Research Genetics (info@resgen.com).
BAC end Web Server: http://www.htsc.washington.edu
Plate: 3122 row: G column: 16
Seq primer: T7
Class: BAC ends
High quality sequence stop: 773.
FEATURES
source
1..773
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Plate=3122 Col=16 Row=G"
/clone_lib="CIT Approved Human Genomic Sperm Library D"
/sex="male"
/note="Organ: sperm; Vector: pBelobAC11; BAC Clones in
E-Coli DH10B"

BASE COUNT      248 a 153 c 84 g 281 t 7 others
ORIGIN
Query Match      19.1%; Score 43; DB 17; Length 773;
Best Local Similarity 59.3%; Pred. No. 22;
Matches 73; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 103 TTTTCAAAGCTACGCTGTCTCCCAACAAACGAACAAACAAACAAACCTTTTAA 162
DB 188 TTTTAAAAAAGGGGCCCTCCCAACAAACGAACAAACAAACAAACCTTTTAA 247

QY 163 GAGTTGATGCTACTCATTTGATCTGCTCTCTCTGATCAATAGGAATTTT 222
DB 248 AAAACCCCTTTTATTTTATTTTTCACAAAAAATTTT 307

QY 223 TTT 225
DB 308 TAT 310

RESULT 6

```

CNS058MA/c  
LOCUS  
DEFINITION  
Tetraodon nigroviridis genome survey sequence T3 end of clone  
002B22 of library A from Tetraodon nigroviridis, genomic survey  
sequence.

ACCESSION  
AL326107.1 GI:8219696  
VERSION  
GSS: genome survey sequence.  
KEYWORDS  
Tetraodon nigroviridis.  
SOURCE  
Tetraodon nigroviridis  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;  
Tetraodontidae; Tetraodon.  
1 (bases 1 to 964)  
Roest-Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,  
Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,  
Saurin,W. and Weissenbach,J.  
Human gene number estimate provided by genome wide analysis using  
Tetraodon nigroviridis DNA sequence  
Unpublished  
2 (bases 1 to 964)  
Roest-Crollius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C.,  
Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and  
Weissenbach,J.  
Characterization and repeat analysis of the compact genome of the  
freshwater pufferfish Tetraodon nigroviridis  
Unpublished  
3 (bases 1 to 964)  
Genoscope.  
Direct Submission  
Submitted (12-APR-2000)  
This sequence is a single read and was generated as part of a large  
scale clone-end sequencing project of the Tetraodon nigroviridis  
genome. For more information, please take a look at  
http://www.genoscope.cns.fr/Tetraodon.  
Location/Qualifiers  
1. 964  
/organism="Tetraodon nigroviridis"  
/db\_xref="taxon:99883"  
/clone="002B22"  
/clone\_lib="A"  
/notes="Genoscope sequence ID : COA002DALL1A1-end : T3"  
BASE COUNT 290 a 194 c 170 g 269 t 41 others  
ORIGIN  
Query Match 18.6%; Score 41.8; DB 17; Length 964;  
Best Local Similarity 52.4%; Pred. No. 34;  
Matches 88; Conservative 0; Mismatches 80; Indels 0; Gaps 0;  
QY 4 TGGTAACTGCTTCAGGCATAAATTGAAATTAATGAGCTCCATCATGCTATA 63  
|||||  
Db 821 TGGTGATTNCCGGTTGTGACCAATTTTAGCCGTACAAATATGGGATTTCTATGTTAA 762  
QY 64 TTGCTTTTACCTTCAGAAAGATATTAGTTTCTACTCAGGTTTTTCAAAGCTAGCGTGCC 123  
|||||  
Db 761 NATTTCTCATTAACACATATGTTTATGTCATGATGATTGTCNAAAATTAGTTGAT 702  
QY 124 CCAAAAACGAAACAAAAACAAACAAACCTTTTAAAGATTGATG 171  
Db 701 TTAATAAAAAACAAAAAATCAAGACAAGTTCTTTTCAGATTCTG 654  
RESULT 7  
CNS00D90/c  
LOCUS  
DEFINITION  
Drosophila melanogaster genome survey sequence TET3 end of BAC #  
BACR27P19 of RPCI-98 library from Drosophila melanogaster (fruit  
fly), genomic survey sequence.  
AL060402  
VERSION  
AL060402.1 GI:4947764  
KEYWORDS  
GSS.  
SOURCE  
Drosophila melanogaster.

CNS058MA  
LOCUS  
DEFINITION  
Tetraodon nigroviridis genome survey sequence T3 end of clone  
002B22 of library A from Tetraodon nigroviridis, genomic survey  
sequence.

ACCESSION  
AL326107.1 GI:8219696  
VERSION  
GSS: genome survey sequence.  
KEYWORDS  
Tetraodon nigroviridis.  
SOURCE  
Tetraodon nigroviridis  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;  
Tetraodontidae; Tetraodon.  
1 (bases 1 to 964)  
Roest-Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,  
Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,  
Saurin,W. and Weissenbach,J.  
Human gene number estimate provided by genome wide analysis using  
Tetraodon nigroviridis DNA sequence  
Unpublished  
2 (bases 1 to 964)  
Roest-Crollius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C.,  
Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and  
Weissenbach,J.  
Characterization and repeat analysis of the compact genome of the  
freshwater pufferfish Tetraodon nigroviridis  
Unpublished  
3 (bases 1 to 964)  
Genoscope.  
Direct Submission  
Submitted (12-APR-2000)  
This sequence is a single read and was generated as part of a large  
scale clone-end sequencing project of the Tetraodon nigroviridis  
genome. For more information, please take a look at  
http://www.genoscope.cns.fr/Tetraodon.  
Location/Qualifiers  
1. 964  
/organism="Tetraodon nigroviridis"  
/db\_xref="taxon:99883"  
/clone="002B22"  
/clone\_lib="A"  
/notes="Genoscope sequence ID : COA002DALL1A1-end : T3"  
BASE COUNT 290 a 194 c 170 g 269 t 41 others  
ORIGIN  
Query Match 18.6%; Score 41.8; DB 17; Length 964;  
Best Local Similarity 52.4%; Pred. No. 34;  
Matches 88; Conservative 0; Mismatches 80; Indels 0; Gaps 0;  
QY 4 TGGTAACTGCTTCAGGCATAAATTGAAATTAATGAGCTCCATCATGCTATA 63  
|||||  
Db 821 TGGTGATTNCCGGTTGTGACCAATTTTAGCCGTACAAATATGGGATTTCTATGTTAA 762  
QY 64 TTGCTTTTACCTTCAGAAAGATATTAGTTTCTACTCAGGTTTTTCAAAGCTAGCGTGCC 123  
|||||  
Db 761 NATTTCTCATTAACACATATGTTTATGTCATGATGATTGTCNAAAATTAGTTGAT 702  
QY 124 CCAAAAACGAAACAAAAACAAACAAACCTTTTAAAGATTGATG 171  
Db 701 TTAATAAAAAACAAAAAATCAAGACAAGTTCTTTTCAGATTCTG 654  
RESULT 7  
CNS00D90/c  
LOCUS  
DEFINITION  
Drosophila melanogaster genome survey sequence TET3 end of BAC #  
BACR27P19 of RPCI-98 library from Drosophila melanogaster (fruit  
fly), genomic survey sequence.  
AL060402  
VERSION  
AL060402.1 GI:4947764  
KEYWORDS  
GSS.  
SOURCE  
Drosophila melanogaster.

Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.  
1 (bases 1 to 1101)  
Genoscope.  
Direct Submission  
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr  
- Web : www.genoscope.cns.fr)  
Determination of this BAC-end sequence was carried out as part of a  
collaboration with the Berkeley Drosophila Genome Project (BDGP).  
The BDGP is constructing a physical map of the Drosophila  
melanogaster genome using these BACs. For further information  
please see http://www.fruitfly.org The BDGP Drosophila  
melanogaster BAC library was prepared by Kazutoyo Osoegawa and  
Aaron Mamoser in Pieter de Jong's laboratory in the Department of  
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,  
NY. The library is named RPCI-98 and was constructed by partial  
EcoRI digestion of Drosophila DNA provided by the BDGP from the  
isogenic strain Y2; cn bw sp, the same strain used for the BDGP's  
P1 and EST libraries. A more detailed description of the library  
and how to order individual BAC clones, the entire library, or  
filters for hybridization from the BACPAC Resource Center can be  
found at http://bacpac.med.buffalo.edu/drosophila\_bac.htm.  
Location/Qualifiers  
1. 1101  
/organism="Drosophila melanogaster"  
/db\_xref="taxon:7227"  
/clone="BACR27P19"  
/clone\_lib="RPCI-98"  
/note="end : TET3"  
BASE COUNT 380 a 163 c 143 g 199 t 216 others  
ORIGIN  
Query Match 18.3%; Score 41.2; DB 17; Length 1101;  
Best Local Similarity 34.5%; Pred. No. 41;  
Matches 58; Conservative 42; Mismatches 68; Indels 0; Gaps 0;  
QY 18 AGCGTAAATTTGAATAAATATGAGCTCCATCATGCTATGCTATGTTTACCTTC 77  
Db 1075 RGAATAAAARADHWNATATATGRTTWWKTTWTTTWTTHTTTAAACWTTTWW 1016  
QY 78 AGAAGATATTTAGTTTCACTCAGGTTTTTCAAAGCTAGCGTGCTCCCAAAACGAAA 137  
Db 1015 AAAAAAAMATTTCTTAAATTTTWWYTCYCWACAMCTCCCTTMCACMAWMCW 956  
QY 138 CAAAAACAAAAACACCTTTTAAAGATTGATGCTACTCATTTGAT 185  
Db 955 CAAAAAAYAAAAATWMAKTTTKTKTCTTWTWTTTTCAMCMAAARWT 908  
RESULT 8  
AW504318 290 bp mRNA linear EST 02-MAR-2000  
LOCUS  
DEFINITION  
UI-HF-BN0-alg-e-05-0-UI-r1 NIH MGC 50 Homo sapiens cDNA clone  
IMAGE:3079401 5', mRNA sequence.  
ACCESSION  
AW504318 GI:7141985  
VERSION  
AW504318.1  
KEYWORDS  
EST.  
SOURCE  
human.  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 290)  
NIH-MGC http://mgi.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-r@mail.nih.gov  
Eco RI site shown at the beginning of the sequence.  
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.  
CDNA Library Preparation: M.B. Soares Lab

CDNA Library Arrayed by: M.B. Soares Lab

DNA Sequencing by: M.B. Soares Lab

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

www-bio-llnl.gov/bbrp/image/image.html

The following repetitive elements were found in this cDNA sequence:

8-63, >POLY A&Simple repeat

Seq primer: M13 Forward.

#### FEATURES

Location/Qualifiers

1..290

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:3079401"

/tissue\_type="lymph"

/cell\_type="germinal center B cells"

/lab\_host="MGC85"

/lab\_host="DH10B (LTI)"

/notes="Vector: p773-Pac; Site 1: NotI; Site 2: Eco RI;

Constructed from size fractionated cytoplasmic mRNA

(3.5-4.4kb). Directionally cloned. Cells provided by

Louis M. Staudt, Ph.D. Library preparation by Maria de

Fatima Bonaldo, Ph.D. and M. Bento Soares, Ph.D."

108 a 51 c 41 g 90 t

BASE COUNT

ORIGIN

Query Match

Best Local Similarity 18.2%; Score 41; DB 10; Length 290;

Matches 74; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

31 GAATAAATATGAGGCTCATATGCTATATGTTTCTCAGAGAAATATTTA 90

111 GAAATAAGTATTTGGGCCCCAAAATGGCATACGGTTTTTCCCAACCTAAGAAGAA 170

91 GTTCTACTCAGGTTTTTCAAAGCTAGCTGCTCCCAAAACGAAACAAACAAA 150

171 GGATGTAATAAATAATTCGGCCCTCCCTCCCAAAACGAAACAAACAAA 230

151 CAACCTTTT 159

231 CAAATTTT 239

RESULT 9

BC022863

LOCUS

DEFINITION

BC022863 Homo sapiens, clone IMAGE:4077047, mRNA.

BC022863.1 GI:22135650

HTC.

WORDS

SOURCE

ORGANISM

Human.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1562)

Strausberg, R.

Submitted (04-FEB-2002) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: [cgabs-r@mail.nih.gov](mailto:cgabs-r@mail.nih.gov)

Tissue Procurement: ATCC

CDNA Library Preparation: CLONTECH Laboratories, Inc.

DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)

<http://www.systemsbio.org>

contact: [amad@systemsbio.org](mailto:amad@systemsbio.org)

Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha

Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

REMARK

COMMENT

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAL Plate: 32 Row: 1 Column: 8

This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein

This clone has the following problem: incomplete processing.

Location/Qualifiers

1..1562

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:4077047"

/tissue\_type="Bone marrow, acute myelogenous leukemia"

/clone\_lib="NIH MGC 55"

/lab\_host="DH10B"

/notes="Vector: pDNR-LIB"

478 a 289 c 331 g 464 t

BASE COUNT

ORIGIN

Query Match

Best Local Similarity 18.2%; Score 41; DB 11; Length 1562;

Matches 83; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

1 TGATGTAGTGTCTTCAGCATAAATTTGAAATAAATATGAGGCTCCATATGCT 60

1407 TGATAATATTTGAAAATGGTGACCAAACTCTATGTAACCTTAGCAAAATTTTCATCT 1466

61 ATATTGGTTTACCTTCAGAAGAAATATTAGTTTCTACTAGGTTTTCRAAGCTACGCTG 120

1467 TTGTGCTTTGGAAATTTAGAAAATATTTAATTTCTTAAAGTCTTATAAAGTTTACTCA 1526

121 TCCCAAAACGAAACCAAAACCAAAACCAAA 153

1527 ATCCCAAAACCAAAACCAAAACCAAAACCA 1559

RESULT 10

AZ523166/c

LOCUS

DEFINITION

AZ523166 216PbB05 Pb MBN #21 Plasmodium berghei genomic 3', DNA sequence.

AZ523166

VERSION

AZ523166.1 GI:13961890

KEYWORDS

GSS.

SOURCE

Plasmodium berghei.

Plasmodium berghei

Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

1 (bases 1 to 595)

Carlton, J.M.-R. and Dame, J.B.

The Plasmodium vivax and P. berghei gene sequence tag projects

Parasitol. Today (Regul. Ed.) 16 (10), 409 (2000)

Contact: Dame JB

Dept. of Pathobiology, College of Veterinary Medicine

University of Florida

2015 SW 23rd Avenue, Bldg 1017, Gainesville, FL 32611, USA

Tel: 352 392 4700

Fax: 352 392 9704

Email: [damej@mail.vetmed.ufl.edu](mailto:damej@mail.vetmed.ufl.edu)

Seq primer: M13(-20) forward

Class: shotgun.

Location/Qualifiers

1..595

/organism="Plasmodium berghei"

/strain="ANKA clone 15cyl (clone of the ANKA 8417 clone)"

/db\_xref="taxon:5821"

/clone\_lib="Pb MBN #21"

/dev\_stage="asexual blood forms"

/lab\_host="Mus musculus"

/note="Vector: pBluescript SK(+)"

excised from lambda ZAP; Site 1: EcoRV; Site 2: EcoRV;

Genomic DNA was prepared from asynchronous blood stage

forms of the cloned ANKA isolate of P. berghei grown in

laboratory Swiss white mice. The DNA was purified from

contaminating host DNA by Hoechst Dye 33258-CsCl ultracentrifugation and precipitated. Purified DNA was digested with mung bean nuclease in the presence of 36-38% formamide at 50 C, as described (Vernick, K.D., Imberski, R.B., and McCutchan, T.F. 1998. Nucleic Acids Research 16:6883-6896). The ends of the digestion fragments were polished using T4 DNA polymerase, and the fragments size selected in the range 500-2000 bp. These were ligated into the EcoRV-cleaved and dephosphorylated pBluescript SK(+) vector. Recombinant plasmids were used to transform E. coli XL10-Gold host cells."

BASE COUNT 236 a 78 c 57 g 224 t  
ORIGIN

Query Match 18.0%; Score 40.6; DB 17; Length 595;  
Best Local Similarity 49.3%; Pred. No. 74;  
Matches 106; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

11 TTGTTTCCAGGCATATAATTTGAATAAATATGAGGCTCCATGATATGCTATATGTTT 70  
569 TTTTTCCTCCCAAAATTTAAATTTTAAATAATATAAATTTCCCAATATGGGAAT 510

71 TACCTTCAGAGAATATTTAGTTTCTACTCAGGTTTTTCAAGCTACGCTGCCCAAAA 130  
509 TTTTTCGAAAAAATAAATTTTTCCTGTTTAAAGAAAAAATTTTCCCTTAAAA 450

131 AACGAAACAAACAAAAACACCTTTTAAAGATTGATGGCTACTCATTTGATCTGCG 190  
449 GGGGATCAAAAAAATAAATTTTTCCTGTTTAAAGAAAAAATTTTCCAGAGGGGCG 390

191 TCCTCTGCTGCAATCAATAGGAATTTTTCCTT 225  
389 ATATTTTTCCTGCAATTCATATGATATATTT 355

RESULT 11  
CNS015W7/c  
LOCUS  
DEFINITION  
Drosophila melanogaster genome survey sequence SP6 end of BAC  
BACN15B11 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.

ACCESSION  
VERSION  
AL105985.1 GI:5619413

KEYWORDS  
SOURCE  
Drosophila melanogaster.

ORGANISM  
Drosophila melanogaster.  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE  
1 (bases 1 to 957)  
Genoscope.

AUTHORS  
TITLE  
JOURNAL

COMMENT  
Direct Submission  
Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)  
Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelOBAC11

FEATURES  
source  
1. .957  
/organism="Drosophila melanogaster"  
/db\_xref="taxon:7227"  
/clone="BACN15B11"  
/clone\_lib="DrosBAC"  
/plasmid="pBelOBAC11"  
/notes="end : SP6"

BASE COUNT 274 a 139 c 139 g 318 t 87 others  
ORIGIN

Query Match 18.0%; Score 40.6; DB 17; Length 957;  
Best Local Similarity 40.1%; Pred. No. 58;  
Matches 73; Conservative 27; Mismatches 82; Indels 0; Gaps 0;

7 TAAAGTTGTTTCCAGGCATATAAATTTGAATAAATATGAGGCTCCATGATATGCTATATG 66  
805 TWATTTTTTTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 746

67 GTTTTACCTTCAGAGAATATTTAGTTTTCACTCAGGTTTTTCAAGCTACGCTGCCCCC 126  
745 WTTTAAATGTTTCAMWAAATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTCGAAG 686

127 AAAAAACCAACCAACCAACCAACCAACCAACCTTTTAAAGAGTTGATGCTACTCATTTGATC 186  
685 WATTATWAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 626

187 TG 188  
625 KK 624

RESULT 12  
AW381028/c  
LOCUS  
DEFINITION  
CM1-HT0285-101199-036-h05 HT0285 Homo sapiens cDNA, mRNA sequence.

ACCESSION  
AW381028

VERSION  
AW381028.1 GI:6885687

KEYWORDS  
SOURCE  
human.

ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE  
1 (bases 1 to 525)  
HCGP http://www.ludwig.org.br/ORESTES.

AUTHORS  
TITLE  
JOURNAL  
COMMENT  
The FAPESP/LICR Human Cancer Genome Project  
Unpublished (1999)  
Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?ti=CM1&t2=CM1-HT0285-101199-036-h05&t3=1999-11-10&t4=1)  
Seq primer: puc 18 forward  
High quality sequence start: 36  
High quality sequence stop: 112.

FEATURES  
Location/Qualifiers  
1. .525  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="HT0285"  
/dev\_stage="Adult"  
/note="Organ: head neck; Vector: puc18; Site: 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196.716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 161 a 97 c 97 g 170 t  
ORIGIN

Query Match 18.0%; Score 40.4; DB 10; Length 525;  
Best Local Similarity 51.7%; Pred. No. 87;  
Matches 92; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

45 GGCTCCATGATATGCTATATGTTTACCTTCAGAGAATAATTTAGTTTCACTCAGGTT 104  
ORIGIN

```

Db      398  GGCCCAAGCTTTACTATATAAGGTTCTGCAAGGTACAAAAGTAAAGTCTGTACAGCT 339
Qy      105  TTTCAAAGCTACGCTGTCCTCCCAAAACGAAACAAACAAACAAACAACTTTTAAAGA 164
Db      338  TTTCAGTAACCTGCTGTCCTCCAGAAAGCAAAATTTAAATATTGCAATCAGCAGATTAAAGGA 279
Qy      165  GTTGATGGCTACTCATTGATCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 222
Db      278  GACTATAGCTAGACATTAGAGCTACAACTTCTGATATCTGTGTGTCAGAAAATCTCTATT 221

RESULT 13
LOCUS      BE325739/c
DEFINITION  NF055G02ST1P1000 Developing stem Medicago truncatula cDNA clone
ACCESSION  BE325739
VERSION    BE325739.2
KEYWORDS   EST.
SOURCE     Bartel medic.
ORGANISM   Medicago truncatula
           Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
           Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
           Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
           Medicago.
REFERENCE  1 (bases 1 to 475)
AUTHORS   He, X.-Z., Shadle, G., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell
           C.J., Flores, H.R., Inman, J.T., Weller, J.W., May, G.D. and Dixon
           R.A.
TITLE     Expressed Sequence Tags from the Samuel Roberts Noble Foundation
JOURNAL   Medicago truncatula stem library
COMMENT   Unpublished (2000)
           On Jul 14, 2000 this sequence version replaced gi:9199516.
           Contact: Dixon RA
           Plant Biology Division
           The Samuel Roberts Noble Foundation
           2510 Sam Noble Parkway, Ardmore, OK 73402, USA
           Tel: 580 221 7302
           Fax: 580 221 7380
           Email: radixon@noble.org
           Medicago Genome Initiative accession: MGI:S:17055
           Insert Length: 685 Std Error: 0.00
           Plate: 055 row: G column: 02
           Seq primer: TCACACAGAAACGCTATGAC.
FEATURES   Location/Qualifiers
            source             1..475
            /organism="Medicago truncatula"
            /db_xref="taxon:3880"
            /clone="NF055G02ST"
            /clone_lib="Developing stem"
            /tissue_type="stem"
            /dev_stage="pooled developmental"
            /notes="Vector: Lambda Zap; Contains a mixture of
            internodal stem segments"
BASE COUNT 182 a 67 c 55 g 171 t
ORIGIN

Query Match 17.8%; Score 40; DB 10; Length 475;
Best Local Similarity 59.9%; Pred. No. 1.1e+02;
Matches 85; Conservative 0; Mismatches 55; Indels 2; Gaps 1;

Qy      11  TTGTTTCAGGCATAAATTTGAATTAATATGAGGCTCCATGATATGCTATATGTTT 70
Db      158  TTGTTTCAGACTTATATTTTGTAGTAATTAATGATT--TTTTTTTGCCAGTTTGGTTG 101
Qy      71  TACCTTCAGAGAATATTTAGTTTCTACTCAGGTTTTTCAAGTACGCTGTCCCAAAA 130
Db      100  TTGCATGTGTGAATAAATACATTTTGATATAGACTATTTCATGCTACTTTTAAAAAAA 41
Qy      131  AACGAAACAAACAAAAACA 152
Db      40  AAAAAAAAAAAAAAAAAAAAA 19

```

```

RESULT 14
LOCUS      AU071524
DEFINITION  AU071524 Dictyostelium discoideum SS (H.Urushihara) Dictyostelium
ACCESSION  AU071524
VERSION    AU071524.1
KEYWORDS   EST.
SOURCE     Dictyostelium discoideum.
ORGANISM   Dictyostelium discoideum.
           Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
REFERENCE  1 (bases 1 to 208)
AUTHORS   Urushihara, H.
TITLE     Developmental cDNA in Dictyostelium discoideum (1999)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Hideko Urushihara
           Institute of Biological Sciences
           University of Tsukuba
           1-1-1 Tennoudai, Tsukuba, Ibaraki 305-8572, Japan
           Tel: 81-298-53-4664
           Fax: 81-298-53-6614
           Email: hideko@biol.tsukuba.ac.jp
           PROJECT = Dictyostelium discoideum cDNA project in Japan.
FEATURES   Location/Qualifiers
            source             1..208
            /organism="Dictyostelium discoideum"
            /strain="AX4"
            /db_xref="taxon:44689"
            /clone="SSB723"
            /clone_lib="Dictyostelium discoideum SS (H.Urushihara)"
            /dev_stage="slug"
BASE COUNT 108 a 11 c 7 g 82 t
ORIGIN

Query Match 17.7%; Score 39.8; DB 9; Length 208;
Best Local Similarity 54.4%; Pred. No. 1.8e+02;
Matches 80; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

Qy      7  TAAGTTGTTTCAGGCATAAATTTGAATTAATATGAGGCTCCATGATATGCTATATTG 66
Db      55  TTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 114
Qy      67  GTTTTACCTTCAGAGAATATTTAGTTTTCACCTCAGGTTTTTCAAAAGCTAGCTGTGCCCC 126
Db      115  TTTTAAATAAAATAAAATTTTATTTATTAACCATAGTTTTTATAGATATATATAGATAA 174
Qy      127  AAAAAACGAAACAAACAAAAACAA 153
Db      175  ACTTATAGCAAAAAAATAAAAAA 201

RESULT 15
LOCUS      B83740/c
DEFINITION  CpG0090B CpIOWAGDNa1 Cryptosporidium parvum genomic, DNA sequence.
ACCESSION  B83740
VERSION    B83740.1
KEYWORDS   GSS.
SOURCE     Cryptosporidium parvum.
ORGANISM   Cryptosporidium parvum.
           Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
           Cryptosporidiidae; Cryptosporidium.
REFERENCE  1 (bases 1 to 647)
AUTHORS   Strong, W.B. and Nelson, R.G.
TITLE     Preliminary profile of the Cryptosporidium parvum genome: an
           expressed sequence tag and genome survey sequence analysis
JOURNAL   Mol. Biochem. Parasitol. 107 (1), 1-32 (2000)
MEDLINE   20183851
COMMENT   Contact: Nelson, R. G.
           Depts. of Medicine & Pharmaceutical Chemistry
           San Francisco General Hospital-University of California, San

```

Francisco  
Box 0811, San Francisco, CA 94143-0811, USA  
Tel: 415 206 8846  
Fax: 415 206 3353  
Email: malari@itsa.ucsf.edu

Submitted sequence has been edited to remove vector sequences 5' to the insert, to correct miscalled bases and assign uncalled (N) bases throughout the sequence, and to terminate when base-calling became ambiguous.

Seq primer: T3

Class: shotgun

High quality sequence stop: 647.

Location/Qualifiers

# FEATURES

source

1. .647  
/organism="Cryptosporidium parvum"  
/strain="IOWA"  
/db\_xref="taxon:5807"  
/clone\_lib="CpIOWAGDNa1"  
/lab\_host="E. coli XL2 Blue MRF"  
/notes="Vector: pBlueScript II (SK-); Site 1: EcoRV; C. parvum (IOWA isolate) genomic DNA was hydrodynamically sheared to produce fragments having a tight size distribution between 2-4 kb by Dr. Yvonne Thorstenson of the Stanford DNA Sequencing and Technology Center (http://sequence-www.stanford.edu/group/techdev/shear.htm)". The randomly sheared gDNA was chromatographed on Sephacryl S-400 to remove any small fragments and DNA eluting in the void volume was subcloned into an EcoR V-digested, alkaline phosphatase-treated pBlueScript II (SK-) vector and transformed into E. coli strain XL2 Blue MRF'. Recombinant clones from the first plating of the library were selected for sequence analysis using T3 and T7 primers."

BASE COUNT 183 a 92 c 90 g 282 t

ORIGIN

Query Match 17.7%; Score 39.8; DB 17; Length 647;  
Best Local Similarity 64.8%; Pred. No. 1e+02;  
Matches 59; Conservative 0; Mismatches 32; Indels 0; Gaps 0;  
Qy 81 AGAATATTAGTTTCACTCAGGTTTTCAGAGCTAGCTGTCCCAAAAAACGAAACAA 140  
Db 630 ATAGATTTCATTCACTAAGTTTGTAGAAAGTCTCTGCTCAAAAAACTCGAACAA 571  
Qy 141 AACAAAAAACACCTTTTAAAGAGTTGATG 171  
570 ATGAAACTATATGAAGATCAGATTGGATG 540

Search completed: June 9, 2003, 10:23:43  
Job time : 1060 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 7, 2003, 14:33:54 ; Search time 154 Seconds  
(without alignments)  
3290.257 Million cell updates/sec

Title: US-09-765-231a-58

Perfect score: 225  
Sequence: 1 ttagtgtagtggttcagg.....attagaattttttttttttt 225

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_101002.\*

- 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*
- 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*
- 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*
- 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*
- 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*
- 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*
- 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*
- 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*
- 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*
- 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*
- 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*
- 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*
- 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*
- 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*
- 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*
- 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*
- 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*
- 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*
- 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*
- 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*
- 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*
- 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*
- 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*
- 24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	225	100.0	225	22	AAH23128 Osteoarthritis tis
2	48.4	21.5	1664976	19	AAV21209 Methanococcus jann
3	41	18.2	5518	24	ABK28306 DNA transcription
4	40	17.8	6809	24	ABN80174 Human chemically m
5	40	17.8	6809	24	ABL70557 Chemically treated
6	40	17.8	6809	24	AA561214 Human gene regulat
7	40	17.8	6809	24	ABK31314 Signal transductio
8	38.2	17.0	640681	24	ABK92287 Buchnera sp. genom
9	38	16.9	12393	24	ABJ33263 Human immune syste

C 10	37.4	16.6	2270	20	AAH23128	Human protein phos
C 11	37.2	16.5	778	22	AAH23128	Human breast cance
C 12	36.6	16.3	2705	19	AAV05164	Human growth arres
C 13	36.4	16.2	345	24	ABQ85902	Arabidopsis thalia
C 14	36.4	16.2	1479	24	ABN95163	Gene #1661 used to
C 15	36.4	16.2	7924	24	ABK40070	Human chemically p
C 16	36.4	16.2	7924	24	ABL34139	Human immune syste
C 17	36.4	16.2	40388	23	ABL26342	Drosophila melanog
C 18	36.2	16.1	5474	24	ABL33270	Human immune syste
C 19	36	16.0	5153	17	AAH230347	Human YAP CDNA. H
C 20	36	16.0	6725	24	ABL33208	Human immune syste
C 21	36	16.0	6725	24	ABL34554	Human metastasis a
C 22	35.8	15.9	502	24	ABQ88918	Human prostate exp
C 23	35.8	15.9	3002	23	ABL15414	Drosophila melanog
C 24	35.8	15.9	7847	24	ABL34188	Human immune syste
C 25	35.6	15.8	6163	24	ABN80119	Human chemically m
C 26	35.4	15.7	932	22	AAK88641	Human digestive sy
C 27	35.4	15.7	113515	24	ABL34175	Human immune syste
C 28	35.2	15.6	557	22	ABL18969	Human breast cance
C 29	35.2	15.6	11049	24	ABL92218	Chemically treated
C 30	35.2	15.6	11049	24	ABL49321	Human polynucleoti
C 31	35.2	15.6	11049	24	ABL32668	Human immune syste
C 32	35	15.6	332	22	AAK55869	Human immune/haema
C 33	35	15.6	1132	23	ABV22830	Human prostate exp
C 34	35	15.6	1132	23	ABV28660	Human prostate exp
C 35	35	15.6	3541	19	AAV07076	CDNA encoding huma
C 36	35	15.6	3541	24	AAH38589	Human stromal cell
C 37	35	15.6	10609	22	AAH46359	Tumour suppressor
C 38	35	15.6	10609	24	ABK31268	Signal transductio
C 39	34.8	15.5	355	22	AAH81841	Rat differential t
C 40	34.8	15.5	388	22	AAH89390	Human polynucleoti
C 41	34.8	15.5	5941	24	ABL33350	Human immune syste
C 42	34.8	15.5	6062	24	ABL34078	Human immune syste
C 43	34.8	15.5	910715	20	AAH20248	Borrelia burgdorfe
C 44	34.6	15.4	5216	24	ABL32347	Human immune syste
C 45	34.6	15.4	5216	24	ABL34459	Human metastasis a

ALIGNMENTS

RESULT 1  
AAH23128  
ID AAH23128 standard; DNA; 225 BP.

XX AAH23128;

XX 17-SEP-2001 (first entry)

XX Osteoarthritis tissue-derived nucleic acid sequence #58.

XX Osteoarthritis; infectious disorder; Crohn's disease; sepsis; human;

XX wound healing; osteopathic; anti-arthritis; anti-inflammatory;

XX vulnery; antibacterial; antiallergic; ds.

XX Homo sapiens.

XX WO200153531-A2.

XX 26-JUL-2001.

XX 18-JAN-2001; 2001WO-US00016.

XX 18-JAN-2000; 2000US-0176523.

XX (PHAA ) PHARMACIA CORP.

XX Phippard D, Vasanthakamur G, Dotson S, Ma X;

XX WPI; 2001-451914/48.

XX Substantially purified protein, polypeptide or their fragments, used to  
PT identify a biologically active compound or composition and treat

mammalian osteoarthritis -

Claim 1; Page 137; 144pp; English.

Sequences AH23071-23152 represent nucleic acid sequences derived from osteoarthritis tissues. The sequences are useful as probes and for the diagnosis or prognosis of mammalian osteoarthritis. The polynucleotides and polypeptides of the invention are useful for generating diagnostic reagents, as targets for small molecule drug development, generation of therapeutics, and cloning genes. Specific antibodies are used to generate enzyme linked immunosorbent assays for detection of osteoarthritis. The invented molecules can be used to treat osteoarthritis or to analyse the disease-modifying activity of osteoarthritis drugs. Other disorders treatable using the nucleic acid sequences include atopic, inflammatory and infectious disorders e.g. Crohn's disease and sepsis, and wound healing.

Sequence 225 BP; 72 A; 38 C; 35 G; 80 T; 0 other;

Query Match 100.0%; Score 225; DB 22; Length 225;

Best Local Similarity 100.0%; Pred. No. 1.2e-45;

Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TGATGGTAAGTGTGTTTCAGGCATATAAATTTGAAATAAATATGAGGCTCCATGATGCT 60

1 TGATGGTAAGTGTGTTTCAGGCATATAAATTTGAAATAAATATGAGGCTCCATGATGCT 60

61 ATATTGTTTTTACCTTCAGAGATATTTAGTTTCTACTCAGGTTTTTCAAGCTACGCTG 120

61 ATATTGTTTTTACCTTCAGAGATATTTAGTTTCTACTCAGGTTTTTCAAGCTACGCTG 120

121 TCCCCCAAAAAACGAAACAAAAACAAAAACAACTTTTTTAAGAGTTGATGGCTACTCAT 180

121 TCCCCCAAAAAACGAAACAAAAACAAAAACAACTTTTTTAAGAGTTGATGGCTACTCAT 180

181 TTGATGTCGCTCTCTGCTGCAATCAATAGGAATTTTTTTTTTTT 225

181 TTGATGTCGCTCTCTGCTGCAATCAATAGGAATTTTTTTTTTTT 225

RESULT 2

AAV21209

ID AAV21209 standard; DNA; 1664976 BP.

XX

AC AAV21209;

XX

DT 10-NOV-1998 (first entry)

Methanococcus jannaschii circular chromosome.

Methanococcus jannaschii; methanogenic archaeon; circular chromosome; genome; autotrophic; extrachromosomal element; identification; ds.

Methanococcus jannaschii.

WO9807830-A2.

26-FEB-1998.

22-AUG-1997; 97WO-US14900.

22-AUG-1996; 96US-0024428.

(GENO-) INST GENOMIC RES.

(UNII) UNIV ILLINOIS FOUND.

(UYJO) UNIV JOHNS HOPKINS SCHOOL MEDICINE.

Bult CJ, Smith HO, Venter JC, White OR, Woese CR;

WPI; 1998-169145/15.

Complete genome sequence of methano-genic archaeon, Methanococcus jannaschii - useful in identification of M. jannaschii genome

fragment

Claim 13; Page 152-585; 614pp; English.

The present sequence represents the complete 1.66-megabase pair genome sequence of the Methanococcus jannaschii circular chromosome. The present invention describes M. jannaschii open reading frames from the genome sequence. The invention also describes a computer based system for identifying fragments of the M. jannaschii genome that are homologous to target nucleotide sequences, comprising: (a) data storage means comprising the nucleotide sequence of the 1664976, 58407 or 16550 bp sequence (see AAV21209, AAV21210 and AAV21211), or a nucleotide target sequence to the nucleotide sequence of the data storage means to identify a homologous sequence, and (c) retrieval means for obtaining the homologous sequence. The method, which is based on whole genome random sequencing of an autotrophic archaeon M. jannaschii, the genome of which consists of 3 physically distinct elements, a large circular chromosome (the 1664976 bp sequence given in AAV21209), a large circular extra-chromosomal element (the 58407 bp sequence given in AAV21210), and a small circular extra-chromosomal element (the 16550 bp sequence given in AAV21211), can be used in the identification of M. jannaschii genome fragment.

Sequence 1664976 BP; 569133 A; 264649 C; 258701 G; 573392 T; 101 other;

Query Match 21.5%; Score 48.4; DB 19; Length 1664976;

Best Local Similarity 51.9%; Pred. No. 0.029;

Matches 109; Conservative 0; Mismatches 101; Indels 0; Gaps 0;

7 TAAGTTGTTTTCAGGCATATAAATTTGAAATAAATATGAGGCTCCATGATGCTATATG 66

151403 TAAATCTCTGAGCCCTCTATTTCATCCATTATTATAGTACATGTTTGTGTAATTTG 151462

67 GTTTTACCTTCAGAGATATTTAGTTTCTACTCAGGTTTTTCAAGCTACGCTGTCCTCC 126

151463 TCTTTATCAATAGCCCTTCTAAAAATCTCTGTAATTTTTTAAACTCCGAGTTGTCT 151522

127 AAAAAACGAAACAAAAACAAAAACAACTTTTTTAAGAGTTGATGGCTACTCATTTGATC 186

151523 AAAATCTATAAAATCAATATTACTCTTTTCCATACTCTTTAATTTGTTTTTATC 151582

187 TGCTCTCTCTGCTGTAATCAATAGGAATTTT 216

151583 TTACCCAAAAATCCCACTTATTAGGAATTT 151612

RESULT 3

ABK28306/C

ID ABK28306 standard; DNA; 5518 BP.

XX

AC ABK28306;

XX

DT 23-APR-2002 (first entry)

XX

DNA transcription associated complementary genomic DNA #90.

DNA transcription associated gene; peptide nucleic acid; PNA-oligomer; PNA; cytosine methylation state; SNP; retroviral infection; gene; ds; single nucleotide polymorphism; adenosine deaminase deficiency; cancer; viral infection; Sezary syndrome; haematological disorder; tuberculosis; immunological disorder; Werner syndrome; developmental disorder; psoriasis; Rieger's syndrome; neurological disorder; erythropoiesis; neurodegenerative disorder; Waardenburg syndrome; Niemann-Pick disease; myelodysplastic syndrome; myocardial infarction; hypertension; arthritis; angiodysplasia; congenital heart disease; HDR syndrome; gene therapy; polyglutamine disorder; solid tumour.

Unidentified.

WO200192565-A2.

06-DEC-2001.

XX 06-APR-2001; 2001WO-EP03973.  
PF XX  
XX 06-APR-2000; 2000DE-1019058.  
PR XX  
PR 07-APR-2000; 2000DE-1019173.  
PR XX  
PR 30-JUN-2000; 2000DE-1032529.  
PR XX  
PR 01-SEP-2000; 2000DE-1043826.  
XX XX  
PA (EPIG-) EPIGENOMICS AG.  
XX  
PI Olek A, Piepenbrock C, Berlin K;  
PI WPI; 2002-050046/12.  
DR XX  
XX New nucleic acids or oligomers, useful for diagnosing or treating  
PT diseases associated with DNA transcription, e.g. immunological  
PT disorders, Werner syndrome, psoriasis, myocardial infarction, solid  
PT tumours or cancer  
XX  
XX Claim 1; SEQ ID No 180; 32pp; English.  
XX  
XX The invention relates to a nucleic acid, which comprises a segment of the  
CC chemically pretreated DNA of genes associated with DNA transcription from  
CC one of 346 sequences, and an oligomer, in particular an oligonucleotide  
CC or peptide nucleic acid (PNA)-oligomer that hybridises to or is identical  
CC to the chemically pretreated DNA of genes associated with DNA  
CC transcription. The set of oligomer probes are useful for detecting the  
CC cytosine methylation state and/or single nucleotide polymorphisms (SNPs)  
CC in a chemically pretreated genomic DNA. The nucleic acids are useful for  
CC diagnosing or treating diseases associated with DNA transcription  
CC (particularly with the methylation status), e.g. adenosine deaminase  
CC deficiency, viral infection, retroviral infection, Sezary syndrome,  
CC haematological disorders, immunological disorders, Werner syndrome,  
CC tuberculosis, developmental disorders, psoriasis, Rieger's syndrome,  
CC neurological disorders, neurodegenerative disorders, Waardenburg  
CC syndrome, Niemann-Pick disease, myelodysplastic syndrome, myocardial  
CC infarction, hypertension, angiogenesis, erythropoiesis, congenital heart  
CC disease, HDR syndrome, arthritis, polyglutamine disorders, solid tumours  
CC or cancer. Sequences ABK28127-ABK28472 represent DNA transcription  
CC associated genomic DNA molecules of the invention.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification but was obtained in electronic format directly from the  
CC European Patent Office.  
XX  
SQ Sequence 5518 BP; 1899 A; 52 C; 1031 G; 2536 T; 0 other;  
Query Match 18.2%; Score 41; DB 24; Length 5518;  
Best Local Similarity 51.4%; Pred. No. 0.87; Indels 0; Gaps 0;  
Matches 95; Conservative 0; Mismatches 90; Indels 0; Gaps 0;  
QY 7 TAAAGTTGTTTCAGGCATATAAATTTGAAATATAATATGAGGCTCCATGATATGCTATATTG 66  
Db 4088 TATATTTTAAACAAAAATTTTCATTAACTATAAACTCTCTAAATATATACGAA 4029  
QY 67 GTTTTACCTTCAGAGAAATATTAGTTTCTACTCAGGTTTTTCAAAGCTAGCTGTCCTCC 126  
Db 4028 TTACTTAAATTAATACTAAATAATAATACTTCAAACTCTTAAAAAATAATACTCTAA 3969.  
QY 127 AAAAAACGAAACAAACAAACAAACCTTTTAAAGTTGATGCTACTCTTTGATC 186  
Db 3968 AATAAAAAATCTTAAATATAAAAAAATACTATTATTTAAATTAATAACAACTTAATCTATT 3909  
QY 187 TGCCT 191  
Db 3908 TAACT 3904  
RESULT 4  
ABN80174/c  
ID ABN80174 standard; DNA; 6809 BP.  
XX  
AC ABN80174;  
XX

DT 15-JUL-2002 (first entry)  
XX Human chemically modified disease associated gene SEQ ID NO 191.  
DE XX  
XX Human; development; homeobox gene; HOX; diabetes; cancer; apoptosis;  
KW heart disease; epilepsy; histone deacetylation; muscular dystrophy;  
KW dwarfism; single nucleotide polymorphism; SNP; cytosine methylation;  
KW antidiabetic; cytostatic; anticonvulsant; ds.  
XX  
XX Homo sapiens.  
OS Synthetic.  
OS  
PN WO200200927-A2.  
XX  
XX 03-JAN-2002.  
PD  
XX 02-JUL-2001; 2001WO-EP07536.  
PF  
XX 30-JUN-2000; 2000DE-1032529.  
PR  
PR 01-SEP-2000; 2000DE-1043826.  
XX  
XX (EPIG-) EPIGENOMICS AG.  
PA  
XX Olek A, Piepenbrock C, Berlin K;  
PI WPI; 2002-130908/17.  
XX  
XX Novel nucleic acid useful for diagnosis and therapy of diseases  
PT associated with development genes such as diabetes, comprises a  
PT sequence of a segment of chemically pretreated DNA of genes associated  
PT with development  
XX  
PS Claim 1; SEQ ID NO 191; 27pp; English.  
XX  
XX The invention relates to a nucleic acid (I) comprising a sequence at  
CC least 18 bases in length of a segment of chemically pretreated DNA (II)  
CC of genes associated with development selected from 87 genes listed in  
CC the specification such as ACCPN, ADFN, or AFD1 and comprising one of 350  
CC sequences (ABN79984-ABN80333) or their complements. The invention is  
CC useful for the diagnosis or therapy of diseases associated with  
CC development genes, in particular disease related to homeobox containing  
CC genes (HOX), like diabetes, cancer, apoptosis related diseases, syndromes  
CC associated with congenital heart disease, epilepsy, diseases related to  
CC histone deacetylation, Currarino syndrome, diseases related with the  
CC development of the brain and limb girdle muscular dystrophy and dwarfism.  
CC Oligomers specific to each of the genes are useful for detecting the  
CC methylation state of all CpG dinucleotides within the 350 sequences or  
CC (II) and their complementary sequences, as primer oligonucleotides for  
CC the amplification of the 350 sequences, (II) and/or their complements and  
CC as oligomer probes for detecting the cytosine methylation state and/or  
CC single nucleotide polymorphisms (SNPs).  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification but is based on sequence information supplied to Derwent by  
CC the European Patent Office.  
XX  
SQ Sequence 6809 BP; 1792 A; 270 C; 1625 G; 3122 T; 0 other;  
Query Match 17.8%; Score 40; DB 24; Length 6809;  
Best Local Similarity 54.9%; Pred. No. 1.6;  
Matches 79; Conservative 0; Mismatches 65; Indels 0; Gaps 0;  
QY 17 CAGGCATAAAATTTGAAATAAATTTATGAGGCTCCATGATATGCTATATTGTTTACCTT 76  
Db 1384 CACAATAAAACCAATATATATCGTAACACACACTCTATTTCATAATTTCTTAAATTAT 1325  
QY 77 CAGAAGAATATTAGTTTCTACTCAGGTTTTTCAAAGCTAGCTGTCCTCCCAAAACGAA 136  
Db 1324 ACAAAACTATTATAAATAAATCCCGTCTATCAATTTTCGTTTACTACAAATTACTTT 1265  
QY 137 ACAAAACAAAAACAAACCTTTT 160  
Db 1264 TCAAACTTAATCAATAAATTCITT 1241

RESULT 5  
ABL70557/c  
ID ABL70557 standard; DNA; 6809 BP.  
XX AC ABL70557;  
XX DT 01-JUL-2002 (first entry)  
XX DE Chemically treated cell signalling DNA sequence#224.  
XX KW Cell signalling; cytosine methylation; cell signalling disease;  
XX KW cancer; tumour; cytostatic; ds.  
XX OS Unidentified.  
XX PN WO200202807-A2.  
XX DT 10-JAN-2002.  
XX PR 29-JUN-2001; 2001WO-EP07471.  
XX PR 30-JUN-2000; 2000DE-1032529.  
XX PR 01-SEP-2000; 2000DE-1043826.  
XX PA (EPIG-) EPIGENOMICS AG.  
XX PI Olek A, Piepenbrock C, Berlin K;  
XX DR WPI; 2002-154758/20.  
XX PT Nucleic acid, useful for diagnosis and therapy of diseases associated  
XX PT with cell signalling e.g. cancer, comprises chemically modified genomic  
XX PT sequences of genes associated with cell signalling  
XX PS Claim 1; SEQ ID NO 447; 24pp-sequence listing; English.  
XX CC The invention relates to a nucleic acid comprising a sequence of at least  
XX CC 18 bases of a segment of chemically pretreated DNA of genes associated  
XX CC with cell signalling. The activity of the modified sequences of the  
XX CC invention may be described as cytostatic. The object of the invention is  
XX CC to provide the chemically modified DNA of genes associated with cell  
XX CC signalling, as well as oligonucleotides and/or PNA-oligomers for  
XX CC detecting cytosine methylations, as well as a method which is  
XX CC particularly suitable for the diagnosis and/or therapy of genetic and  
XX CC epigenetic parameters of genes associated with cell signalling. The  
XX CC chemically modified DNA provided by the invention is useful for diagnosis  
XX CC and therapy of diseases such as solid tumours and cancer. The sequences  
XX CC given in records ABL70111-ABL70626 represent chemically pre-treated  
XX CC genomic DNA's of genes associated with cell signalling.  
XX CC Note: The sequence data for this patent is not represented in the printed  
XX CC specification, but is based on sequence information supplied by the  
XX CC European Patent Office.  
XX SQ Sequence 6809 BP; 1792 A; 270 C; 1625 G; 3122 T; 0 other;  
Query Match 17.8%; Score 40; DB 24; Length 6809;  
Best Local Similarity 54.9%; Pred. No. 1.6;  
Matches 79; Conservative 0; Mismatches 65; Indels 0; Gaps 0;  
QY 17 CAGGCATATAATTTGAAATTAATTTAGAGCTCCATGATGCTATATGTTTACCTT 76  
DB 1384 CACAATAAACCACATATATATCGTAACACACACTCTATTCTTAATTTAT 1325  
QY 77 CAGAAGATATTTAGTTTCTACTCAGGTTTTTCAAAGCTACGCTGCCCAAAAACGAA 136  
DB 1324 ACAAAACTATTTAAATTAATTAATCCGCTCTATCAATTTTCGTTTACTACAAATTTACTTT 1265  
QY 137 ACAAAACAAAACACACCTTTT 160  
DB 1264 TCAAAACTTAATCAATAATTTCTTT 1241

RESULT 6  
AAS61214/c  
ID AAS61214 standard; DNA; 6809 BP.  
XX AC AAS61214;  
XX DT 29-JAN-2002 (first entry)  
XX DE Human gene regulation-associated gene oligonucleotide #169.  
XX KW Human; Gene regulation-associated gene; severe combined immunodeficiency;  
XX KW cardiac damage; inflammatory response; haemophilia; Werner syndrome;  
XX KW asthma; HDR syndrome; congenital heart defect; Saethre-Chotzen syndrome;  
XX KW renal disease; Preclampsia; cardiac allograft vascular disease;  
XX KW colorectal cancer; thyroid cancer; oesophageal cancer; ds; tumour;  
XX KW immunostimulant; cardiant; antiinflammatory; coagulant; antiasthmatic;  
XX KW nephrotropic; gynecological; anti-tumour; immunosuppressive; cytostatic.  
XX OS Homo sapiens.  
XX PN WO200177375-A2.  
XX PD 18-OCT-2001.  
XX PF 06-APR-2001; 2001WO-EP03968.  
XX PR 06-APR-2000; 2000DE-1019058.  
XX PR 07-APR-2000; 2000DE-1019173.  
XX PR 30-JUN-2000; 2000DE-1032529.  
XX PR 01-SEP-2000; 2000DE-1043826.  
XX PA (EPIG-) EPIGENOMICS AG.  
XX PI Olek A, Piepenbrock C, Berlin K;  
XX DR WPI; 2002-017470/02.  
XX PT New nucleic acid sequences from chemically modified genes associated  
XX PT with gene regulation, useful for analysing cytosine methylations for  
XX PT diagnosis and therapy of diseases e.g. severe combined immunodeficiency  
XX PT disease  
XX PS Claim 1; SEQ ID No 173; 26pp; English.  
XX CC The invention relates to 224 nucleic acid sequences comprising at least  
XX CC 18 bases of a chemically pretreated gene associated with gene regulation  
XX CC selected from 43 known genes (or complementary sequences). The  
XX CC chemical pretreatment converts cytosine bases unmethylated at the  
XX CC 5-position to uracil or another base with hybridisation behaviour  
XX CC dissimilar to cytosine, to enable analysis of cytosine methylations.  
XX CC The DNA sequences, oligomers (or sets/arrays) and method are  
XX CC useful in the diagnosis of diseases (or predisposition to diseases)  
XX CC associated with gene regulation and in therapy of such diseases, by  
XX CC enabling analysis of the cytosine methylation patterns of such genes,  
XX CC kits are provided. They are especially useful in diagnosis  
XX CC and therapy of e.g. severe combined immunodeficiency disease, cardiac  
XX CC disorders, haemophilia, solid tumours and cancer, Werner syndrome,  
XX CC asthma, HDR syndrome, Saethre-Chotzen syndrome, renal disease,  
XX CC preclampsia, graft versus-host disease. The present sequence is a  
XX CC sequence included in the sequence data for this specification and is  
XX CC associated with the human gene regulation-associated genes.  
XX CC Note: The sequence data for this patent did not form part  
XX CC of the printed specification, but was obtained in electronic  
XX CC format directly from WIPO at  
XX CC ftp.wipo.int/pub/published\_pct\_sequences  
XX SQ Sequence 6809 BP; 1792 A; 270 C; 1625 G; 3122 T; 0 other;  
Query Match 17.8%; Score 40; DB 24; Length 6809;  
Best Local Similarity 54.9%; Pred. No. 1.6;  
Matches 79; Conservative 0; Mismatches 65; Indels 0; Gaps 0;  
QY 17 CAGGCATATAATTTGAAATTAATTTAGAGCTCCATGATGCTATATGTTTACCTT 76

Db 1384 CACAATATAAACAATATATATATGTAACACACACTCTATTCATATTTCTTTAATAT 1325  
QY 77 CAGAAGATATTTAGTTTCACTCAGGTTTTTCAAGCTCAGGCTGCTCCCAACAAACGAA 136  
Db 1324 ACAAAACTATTTAAATTAATCCGCTCATCAATTTTCGTTTACTACAAATTAATTT 1265  
QY 137 ACAAAACAAAACAAACCTTTT 160  
Db 1264 TCAAACTTAATCATAAATTTCTTT 1241

## RESULT 7

ABK31314/c  
ID ABK31314 standard; DNA; 6809 BP.

XX AC ABK31314;

DT 23-APR-2002 (first entry)

XX Signal transduction associated gene modified DNA #79.

KW Human; signal transduction associated gene; cytosine methylation state;  
KW CpG island; signal transduction associated disease; solid tumour; cancer;  
KW antitumour; cytostatic; mutant; ds.

XX Homo sapiens.

OS Synthetic.

PN WO200200926-A2.

XX 03-JAN-2002.

XX 29-JUN-2001; 2001WO-EP07472.

XX 30-JUN-2000; 2000DE-1032529.

PR 01-SEP-2000; 2000DE-1043826.

XX (SPIG-) EPIGENOMICS AG.

XX Olek A, Piepenbrock C, Berlin K;

XX WPI; 2002-147896/19.

XX Oligonucleotide for diagnosis and therapy of diseases associated with  
PT signal transduction e.g. cancer, comprises chemically modified genomic  
PT sequences of genes associated with signal transduction

XX Claim 1; SEQ ID No 157; 24pp; English.

CC The present invention relates to chemically modified DNA sequences of  
CC signal transduction associated genes. The DNA sequences are chemically  
CC modified using a solution of bisulphite, hydrogen sulphite or  
CC disulphite. Also disclosed are oligonucleotides and/or PNA oligomers  
CC for detecting the cytosine methylation state (CpG islands) of these  
CC genes, and a method for the diagnosis and/or therapy of genetic and  
CC epigenetic parameters of genes associated with signal transduction.  
CC The genomic DNA can be obtained from cells or cellular components which  
CC contain DNA, e.g. cell lines, biopsies, blood, sputum, stool, urine,  
CC cerebral-spinal fluid, tissue embedded in paraffin such as tissue from  
CC eyes, intestine, kidney, brain, heart, prostate, lung, breast or liver,  
CC histologic object slides, and all their possible combinations. The  
CC sequences of the invention are useful for the diagnosis and therapy of  
CC diseases associated with signal transduction e.g. solid tumours and  
CC cancer. ABK31158-ABK31545 represent chemically pretreated genomic DNA  
CC sequences of different genes associated with signal transduction, or  
CC their complementary sequences.

CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from the  
CC European Patent Office.

XX Sequence 6809 BP; 1792 A; 270 C; 1625 G; 3122 T; 0 other;

Query Match 17.8%; Score 40; DB 24; Length 6809;  
Best Local Similarity 54.9%; Pred. No. 1.6;  
Matches 79; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY 17 CAGGCATATAATTTGAAATAAATATGAGGCTCCATGATATGCTATATGTTTACCTT 76

Db 1384 CACAATAAACAATATATATATGTAACACACACTCTATTCATATTTCTTTAATAT 1325

QY 77 CAGAAGATATTTAGTTTCACTCAGGTTTTTCAAGCTCAGGCTGCTCCCAACAAACGAA 136

Db 1324 ACAAAACTATTTAAATTAATCCGCTCATCAATTTTCGTTTACTACAAATTAATTT 1265

QY 137 ACAAAACAAAACAAACCTTTT 160

Db 1264 TCAAACTTAATCATAAATTTCTTT 1241

## RESULT 8

ABA92787/c

ID ABA92787 standard; DNA; 640681 BP.

XX AC ABA92787;

DT 27-MAR-2002 (first entry)

XX Buchnera sp. genomic DNA SEQ ID NO:1.

DE Buchnera; cockroach-symbiotic bacterium; cockroach extermination;  
KW circular; ds.

XX Buchnera sp.

XX JP2001292771-A.

XX 23-OCT-2001.

XX 07-APR-2000; 2000JP-0107160.

XX 07-APR-2000; 2000JP-0107160.

XX (RIKA) RIKAGAKU KENKYUSHO.

XX WPI; 2002-126043/17.

XX A genomic DNA of cockroach-symbiotic bacterium

XX Claim 1; Page 16-230; 237pp; Japanese.

CC The present invention describes a gene (I) derived from Buchnera sp.  
CC containing the DNA (a) or (b), (a) has a fully defined base pair  
CC sequence selected from a table of sequences found in the Buchnera sp.  
CC genomic DNA of ABA92787 given in the specification or is a DNA selected  
CC from complementary DNA sequences, and (b) is a DNA which hybridises with  
CC the DNA (a) and encodes a protein. Also described are: (1) a recombinant  
CC vector (II) containing (I); (2) a transformant (III) containing (II);  
CC (3) a genomic DNA of Buchnera sp. containing the sequence given in  
CC ABA92787; (4) a plasmid derived from Buchnera sp. containing (c) or  
CC (d), (c) is a DNA containing a fully defined sequence given in ABA92788  
CC or ABA92789 and (d) is a plasmid which hybridises with a DNA; and (5) a  
CC method for the preparation of a protein in which (III) is cultured and  
CC the expression protein of the objective protein is collected from the  
CC resultant culture. The DNA is useful for developing agricultural  
CC chemicals for exterminating cockroaches. The present sequence represents  
CC the specifically claimed Buchnera sp. genomic DNA sequence, from the  
CC present invention.

XX Sequence 640681 BP; 237522 A; 83822 C; 84757 G; 234580 T; 0 other;

Query Match 17.0%; Score 38.2; DB 24; Length 640681;

Best Local Similarity 55.7%; Pred. No. 7.7;

Matches 73; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 32 AAATAAATTTATGAGGCTCCATGATATGCTATATTTGTTTACCTTCAGAAGATATTTAG 91

Db 198302 AATCAAAATTAATAAGTAGTATATTTTCAATTATGACTTTTCAATAATTTTATTATTAAAG 198243  
QY 92 TTTCACCTCAGGTTTTTCAAGCTACGCTGTCCTCCCAAAACGAAACAAAACAAAAC 151  
Db 198242 TAATAATTTGTTATTATTAATAATACTTTTCATCTAAATAAACAATACAAAACAAAACAAA 198183  
QY 152 AACCTTTTAA 162  
Db 198182 AAGATTTTAA 198172

RESULT 9  
ABL33263/c  
ID ABL33263 standard; DNA; 12393 BP.  
XX  
AC ABL33263;  
XX  
DT 26-MAR-2002 (first entry)  
XX Human immune system associated gene SEQ ID NO: 1236.  
XX Human; immune system disease; cytosine methylation; antiasthmatic;  
KW antiarteriosclerotic; antianemic; cytosine; cytosine; cytosine;  
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;  
KW antiinflammatory; antiarthritic; antidiabetic; antipsoriatic;  
KW acute myeloid leukaemia; cancer; eye disease; arteriosclerosis; anaemia;  
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;  
KW gene; ds.  
XX  
OS Homo sapiens.  
XX  
PN W0200200928-A2.  
XX  
PD 03-JAN-2002.  
XX  
PF 02-JUL-2001; 2001MO-EP07537.  
XX  
PR 30-JUN-2000; 2000DE-1032529.  
XX  
PR 01-SEP-2000; 2000DE-1043826.  
XX  
PA (EPIG-) EPIGENOMICS AG.  
XX  
PI Olek A, Piepenbrock C, Berlin K;  
XX  
XX WPI; 2002-130909/17.

Nucleic acid comprising fragment of chemically modified gene, useful for diagnosis and treatment of diseases associated with abnormal cytosine methylation -

Claim 1; SEQ ID NO 1236; 32pp + Sequence Listing; German.

The present invention provides a number of human immune system associated genes which are modified by the methylation of cytosines. The sequences can be used in the diagnosis and treatment of immune system disorders, including eye diseases such as retinopathy, neovascular glaucoma and macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis, rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel diseases. The present sequence is a gene of the invention.

Sequence 12393 BP; 3484 A; 219 C; 2406 G; 6282 T; 2 other;

Query Match 16.9%; Score 38; DB 24; Length 12393;  
Best Local Similarity 49.5%; Pred. No. 5.2;  
Matches 98; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

QY 22 ATAAATTTGAATAATTAAGCTCCATGATATCTATATGTTTACCTTCAGAA 81  
Db 6387 AAAAAAATTAATAATTTTAAATAATTTTAAATAATTTTCTCTCTTAA 6328

QY 82 GAATATTAGTTTCACTCAGGTTTTTCAAGCTACGCTGTCCTCCCAAAACGAAACAAA 141  
Db 6327 AAAACATAATTTTATTAAACGATTTTAAAAAATCAAAATACTCGATAAATTAATAAA 6268  
QY 142 AAAAAAACAACCTTTTAAAGAGTTGATGGCTACTCATTTGATCGCTCTCTCTGCTGA 201  
Db 6267 TAAATAAAAAATCAAAATATCTCAATTTCTCCCGCTCTCTCAAAATAACCTTA 6208  
QY 202 ATCAATTAGGAATTTT 219  
Db 6207 AACAAATTTCTATCATTT 6190

RESULT 10  
AAX05715/c  
ID AAX05715 standard; DNA; 2270 BP.  
XX  
AC AAX05715;  
XX  
DT 07-MAY-1999 (first entry)  
XX Human protein phosphatase (PROPHO) encoding DNA.  
XX  
DE Human protein phosphatase (PROPHO) encoding DNA.  
XX  
KW Protein phosphatase; PROPHO; apoptosis; AIDS; Alzheimer's Disease;  
KW Acquired Immune Deficiency Syndrome; Parkinson's Disease; inflammation;  
KW cell proliferation; Addison's disease; allergy; anemia; cancer; bone;  
KW leukemia; breast; brain; human; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key  
FT 70..1509  
CDS /tag= a  
FT /product= "protein phosphatase (PROPHO)"  
XX

PN W09856925-A1.

XX 17-DEC-1998.

XX 11-JUN-1998; 98WO-US11614.

XX 11-JUN-1997; 97US-0873093.

XX (INCY-) INCYTE PHARM INC.

XX Bandman O, Corley NC, Goli SK, Lal P, Zhang H;

XX WPI; 1999-080906/07.

XX P-PSDB; AAW94283.

XX New substantially purified human protein phosphatase (PROPHO) -  
XX useful in the diagnosis, prevention or treatment of inflammation,  
XX cancer, and disorders associated with apoptosis

XX Claim 5; Fig 1A-G; 73pp; English.

XX This DNA encodes a human protein phosphatase (PROPHO). Host cells  
XX containing a vector comprising the PROPHO nucleic acid are used for the  
XX recombinant production of the protein. PROPHO forms a composition in the  
XX treatment or prevention of apoptosis-related disorders (e.g. Acquired  
XX Immune Deficiency Syndrome (AIDS), Alzheimer's Disease and Parkinson's  
XX Disease), and in the stimulation of cell proliferation. Antagonists of  
XX the protein are useful in treating inflammation (e.g. Addison's disease,  
XX allergies and anemia), and disorders associated with cell proliferation  
XX (including various cancers like leukemia, and cancers affecting bone,  
XX breast and brain). Complementary polynucleotides are useful in detecting  
XX polynucleotides that encode PROPHO, useful in the diagnosis of conditions  
XX associated with the expression of PROPHO, and in assays that detect  
XX activation or induction of various cancers. PROPHO is useful in producing  
XX antibodies or screening libraries of pharmaceutical agents in order to  
XX identify those that bind to PROPHO.

XX Sequence 2270 BP; 693 A; 344 C; 515 G; 716 T; 2 other;

Query Match 15.6%; Score 37.4; DB 20; Length 2270;  
Best Local Similarity 53.0%; Pred. No. 5.8;  
Matches 80; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 29 TTGAATAAATATGAGGCTCCATGATATGCTATATTGGTTTACCTTCAGAGAATATT 88  
DB 2197 TTGCATACATTTTGATGTGTGTAGTTAGGCTATACAGATATACAAACCAAGCTGTAT 2138

QY 89 TAGTTTCACCTCAGGTTTTCAAGCTACGCTGTCCCAAAACGAAACAAACAAAAA 148  
DB 2137 GAATAATACATAGGTTTTCAAAAGTTATGTTTCATAAAGATACAGAGAAGGAAAAAT 2078

QY 149 AACCAACCTTTTAAGAGTTGATGCTACTCA 179  
DB 2077 AGCCACACATCCAATATCTCACACCTTCTAA 2047

## RESULT 11

AA121862/C  
AA121862 standard; cDNA; 778 BP.

AA121862;  
07-DEC-2001 (first entry)  
Human breast cancer expressed polynucleotide 14319.  
Human; breast cancer; cell marker; cytostatic; ss.  
Homo sapiens.

WO200151628-A2.

19-JUL-2001.

10-JAN-2001; 2001WO-US00798.

14-JAN-2000; 2000US-0176077.

14-MAR-2000; 2000US-0189167.

24-MAR-2000; 2000US-0192099.

29-MAR-2000; 2000US-0193480.

15-MAY-2000; 2000US-0205230.

09-JUN-2000; 2000US-0211315.

25-JUL-2000; 2000US-0220534.

(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

Lillie J, Xu Y, Wang Y, Steinmann K;

WPI; 2001-451856/48.

New peptide useful as a marker for the diagnosis of breast cancer -

Claim 1; Page 2564-2565; 3695pp; English.

The invention relates to human breast cancer expressed polynucleotides (AA107544-AA126789) and methods of assessing whether a patient is afflicted with breast cancer by examining the correlation between the expression of certain markers and the cancerous state of breast cells. The polynucleotides and encoded polypeptides are potential markers for detecting, diagnosing, monitoring, characterising treating and potentially preventing breast cancer. The polynucleotides and encoded polypeptides are also useful for isolating compounds with cytostatic activity.

Sequence 778 BP; 220 A; 134 C; 180 G; 244 T; 0 other;

Query Match 16.5%; Score 37.2; DB 22; Length 778;  
Best Local Similarity 55.4%; Pred. No. 5.7;  
Matches 72; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 27 ATTTGAATAAATATGAGGCTCCATGATATGCTATATTGGTTTACCTTCAGAGAATA 86

Db 136 AATTTAGCTATATTATCTTCTCAATCAAAATTTACTTTCAGAGTAAGTAAAAAAGAGT 77  
QY 87 TTTAGTTTCACTCAGGTTTTTCAAAGCTACGCTGTCCCAAAACGAAACAAACAAA 146  
Db 76 TATAGCAATACATATAATGGACAAAAGGTATCCCAAAAAAAGAAAAA 17  
QY 147 AAAACAACCT 156  
Db 16 AAAAGTACCT 7

## RESULT 12

AAV05164  
ID AAV05164 standard; cDNA; 2705 BP.

XX AC AAV05164;

XX DT 06-JUL-1998 (first entry)

XX DE Human growth arrest gene B4B.

XX DE Growth arrest gene; B4B gene; cell proliferation; cancer; leukaemia;  
XX KW marker; gene therapy; human; ss.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX CDS 123..596

XX FT /\*tag= "a"

XX FT /note= "(Claim 4)"

XX PN WO9800544-A2.

XX PD 08-JAN-1998.

XX PF 27-JUN-1997; 97WO-US11341.

XX PR 28-JUN-1996; 96US-0021874.

XX PA (ACTI-) ACTIVATED CELL THERAPY INC.

XX PI Engleman EG, Laus R, Ruegg CL;

XX DR WPI; 1998-086969/08.

XX DR P-PSDB; AAW46489.

XX PT New isolated DNA encoding growth arresting protein expressed in B  
XX PT cells - and related vectors, transformed cells and proteins, used  
XX PT for, e.g. gene therapy of leukaemia or diagnosis of cancer and  
XX PT immune status

XX PS Claim 4; Fig 1; 26pp; English.

XX CC This cDNA clone, the coding region of which is claimed, codes for  
XX CC a growth arrest protein (see AAW46489). It corresponds to the B4B  
XX CC gene, expression of which results in inhibition of cellular  
XX CC proliferation. Human peripheral blood mononuclear cells were  
XX CC fractionated to recover high, intermediate and low density  
XX CC fractions. cDNA libraries from all 3 fractions were subjected to  
XX CC a differential display PCR. Bands unique to the intermediate  
XX CC density sample were reamplified, cloned and sequenced. A 101 bp  
XX CC insert was used to screen a cDNA library from intermediate  
XX CC density cells to isolate the 2.7 kb cDNA clone. This insert was  
XX CC used to isolate a genomic fragment localised to the B4B gene to  
XX CC chromosome 20q12-q13.1. The isolated clone can be used to express  
XX CC the growth arrest protein in host cells. Introduction of a  
XX CC functional B4B allele (by gene therapy) into a cell with a  
XX CC defective allele may be used to prevent development of pro-B-cell  
XX CC leukaemia or other malignancies. Introduction of the isolated DNA  
XX CC in a vector that includes a suicide inhibition cassette can be used  
XX CC to arrest expansion of cells, in vivo or in vitro, e.g. useful in  
XX CC skin cell grafts and in lymphocytes containing a heterologous

CC cytokine gene for in vivo expression.

XX Sequence 2705 BP; 729 A; 584 C; 627 G; 765 T; 0 other;

SQ Query Match 16.3%; Score 36.6; DB 19; Length 2705;

Best Local Similarity 57.4%; Pred. No. 9.3;

Matches 66; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY. 39 TTATGAGCTCCATGATATCTATATTTGTTTACCTTCAGAGAATATTAGTTTCACT 98

DB 2588 TTAATAGATTATTTTCATATCTATATTTGTTTAAATATTTTGCATACAAATGTTTATACTCT 2647

QY 99 CAGGTTTTTCAAGCTACGCTGTCCTCCCAAAAACGAAACAAAACAAAACAA 153

DB 2648 AGGATATAAAACAGATTCTGATTCCTCCCAAAAACAAAACAAAACAAAAC 2702

RESULT 13

ABQ85902/c

ABQ85902 standard; DNA; 345 BP.

ABQ85902;

05-SEP-2002 (first entry)

Arabidopsis thaliana expressed polynucleotide SEQ ID NO 772.

Plant; Arabidopsis; transgenic; fungicide; insecticide; ds.

Arabidopsis thaliana.

US2002062014-A1.

23-MAY-2002.

26-JAN-2001; 2001US-0770791.

27-JAN-2000; 2000US-178480P.

(GOL/) GORLACH J.

(ANY/) AN Y.

(HAMI/) HAMILTON C M.

(PRIC/) PRICE J L.

(RAIN/) RAINES T M.

(YUY/) YU Y.

(RAME/) RAMEKA J G.

(PAGE/) PAGE A.

(MATH/) MATHW A V.

(LEDF/) LEDFORD B L.

(WOES/) WOESSNER J P.

(HAAS/) HAAS W D.

(GARC/) GARCIA C A.

(KRIC/) KRICKER M.

(SLAT/) SLATER T.

(DAVI/) DAVIS K R.

(ALLE/) ALLEN K.

(HOFF/) HOFFMAN N.

(HURB/) HURBAN P.

Gorlach J, An Y, Hamilton CM, Price JL, Raines TM, Yu Y;

PI Rameka JG, Page A, Mathew AV, Ledford BL, Woessner JP, Haas WD;

PI Garcia CA, Kricker M, Slater T, Davis KR, Allen K, Hoffman N;

PI Hurban P;

XX WPI; 2002-479265/51.

New nucleic acid sequences of Arabidopsis thaliana and their encoded

products are useful to produce transgenic plants, to screen for

biologically active agents such as fungicides and insecticides and in

genetic studies

XX Claim 1; SEQ ID NO 772; 18pp + Sequence Listing; English.

CC

CC The invention relates to a novel nucleic acid of Arabidopsis thaliana

comprising a sequence capable of hybridizing under stringency to one of

the 999 sequences referred to but not defined in the specification

(AB08131-AB086129). The nucleic acid sequences are useful to identify

CC homologous or related genes, to produce compositions that modulate

CC expression or function of the encoded protein, to map functional regions

CC of the protein, to study associated physiological pathways, to

CC genetically manipulate cells and plants. The encoded products are useful

CC to screen for biologically active agents such as fungicides or

CC insecticides and to elucidate biochemical pathways.

XX Sequence 345 BP; 97 A; 61 C; 48 G; 139 T; 0 other;

SQ Query Match 16.2%; Score 36.4; DB 24; Length 345;

Best Local Similarity 53.5%; Pred. No. 8;

Matches 76; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY 12 TGTTCAGGCATAAAATTTGAAATAAATATATGAGGCTCCATATGCTATATTTGTTTT 71

DB 155 TGTATTACCTTAAGACAATAAATATGATTTATGAATAAAATAATTTGAATTTACTGCATAT 96

QY 72 ACCTTCAGAGAATATTTAGTTTCACTCAGGTTTTTCAAGCTACGCTGCTCCCAAAAA 131

DB 95 GAAATATATAATCTTTTGTGTTTAACTCAAGATGTAATAATTTATGATCTTCACATAAA 36

QY 132 ACGAAACAAACAAAAACAA 153

DB 35 AAAAAAAAAAAAAAAAAAAAAA 14

RESULT 14

ABN95163

ID ABN95163 standard; DNA; 1479 BP.

AC ABN95163;

DT 13-AUG-2002 (first entry)

XX Gene #1661 used to diagnose liver cancer.

XX Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;

KW metastatic liver tumour; cytostatic; expression profile; disease state;

KW disease progression; drug toxicity; drug efficacy; drug metabolism.

OS Homo sapiens.

XX WO200229103-A2.

XX 11-APR-2002.

XX 02-OCT-2001; 2001WO-US30589.

XX 02-OCT-2000; 2000US-237054P.

(GENE-) GENE LOGIC INC.

Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;

WPI; 2002-426119/45.

Diagnosing and detecting the progression of liver cancer,

PT hepatocellular carcinoma or metastatic liver tumor in a patient,

PT involves detecting the level of expression of two or more genes in a

PT liver tissue sample

XX Claim 1; SEQ ID NO 1661; 298pp; English.

The invention relates to a novel method for diagnosing and detecting the

progression of liver cancer, hepatocellular carcinoma or metastatic liver

tumor in a patient, and differentiating metastatic liver cancer from

CC hepatocellular carcinoma in a patient, involving detecting the level of

expression of two or more genes represented in ABN93503-ABN97455 in a

tissue sample. The method of the invention has hepatotropic, and



CC cytostatic activity. The method is useful for diagnosing and detecting  
CC the progression of liver cancer, hepatocellular carcinoma and metastatic  
CC liver carcinoma in a patient. The method is useful for identifying  
CC expression profiles which serve as useful diagnostic markers as well as  
CC markers that can be used to monitor disease states, disease progression,  
CC drug toxicity, drug efficacy and drug metabolism.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX  
SQ Sequence 1479 BP; 432 A; 287 C; 266 G; 493 T; 1 other;

Query Match 16.2%; Score 36.4; DB 24; Length 1479;

Best Local Similarity 61.7%; Pred. No. 9.6; Mismatches 0; Indels 0; Gaps 0;  
Matches 58; Conservative 0;

OY 128 AAAAAGCAACAAACAAACAAACAAACCTTTTATAGAGTTGATGGTACTACTTTGATCT 187

DB 834 AGAAGAAGCAACAAACAAACAAACCTTTTATAGAGTTGATGGTACTACTTTGATCT 893

DB 188 GCCTCTCTGCTGAATCAATTAGGAATTTTTTTT 221

DB 894 CCCTGCGTACTGCTCTTTCCAGGAACCTTTACTT 927

#### RESULT 15

ABK40070/c

ID ABK40070 standard; DNA; 7924 BP.

XX AC ABK40070;

XX DT 21-MAY-2002 (first entry)

XX DE Human chemically pretreated gene sequence #76 strand 2.

XX KW Human; ds; bisulphite treatment; CpG; DNA methylation; cancer; tumour;

XX KW Cytotoxic; ALDH6; CYP11A; CYP11B; CYP3A3; DPYD; EPHX2; OCLN; TXNRD1;

XX KW UGT8; MRP; pharmacogenomics; SNP; single nucleotide polymorphism.

XX OS Homo sapiens.

XX PN WO200202806-A2.

XX PD 10-JAN-2002.

XX PF 29-JUN-2001; 2001WO-EP07470.

XX PR 30-JUN-2000; 2000DE-1032529.

XX PR 01-SEP-2000; 2000DE-1043826.

XX (EPIC-) EPIGENOMICS AG.

XX PI Olek A, Piepenbrock C, Berlin K;

XX DR WPI; 2002-154757/20.

XX New nucleic acid, oligonucleotides and peptide nucleic acid-oligomers,

XX PT useful for detecting cytosine methylation state of genes associated

XX PT with pharmacogenomics and for therapy of diseases e.g. cancer

XX PS Claim 1; SEQ ID No 152; 24pp; English.

XX The invention relates to a nucleic acid comprising a sequence at

XX least 18 bases in length of a segment of the chemically pretreated DNA

XX of genes associated with pharmacogenomics according to one of the

XX sequences of the genes ALDH6 (NM\_000693), CYP11A (NM\_000781), CYP11B1

XX (NM\_000497), CYP3A3 (NM\_000776 and NM\_017460), DPYD (NM\_000110), EPHX2

XX (NM\_001979), OCLN (NM\_002538), TXNRD1 (NM\_003330), UGT8 (NM\_003360),

XX MRP (NM\_004996), NM\_019900, NM\_019901, NM\_019902, NM\_019862, NM\_019898,

XX NM\_019859) and their complementary sequences, or a sequence (S1) chosen

XX from 87 sequences and their complements. The chemical pretreatment

XX is bisulphite treatment to convert cytosines (but not methyl-cytosines)

XX into uracils. Also included are an oligomer (II) in particular an

CC oligonucleotide or a peptide nucleic acid (PNA)-oligomer, comprising in  
CC each case at least one base sequence having a length of 9 nucleotides  
CC which hybridises to or is identical to a chemically pretreated DNA of  
CC genes associated with pharmacogenomics and their complements, arranged in  
CC an array for analysing diseases associated with the methylation state  
CC (CpG) and/or detecting SNPs (single nucleotide polymorphisms)  
CC of the 87 sequences. The oligomers may also be used as PCR primers.  
CC The set of 87 nucleic acids and their complements is useful for diagnosis  
CC and therapy of solid tumours and cancer. The present sequence  
CC represents one the 87 DNA sequences or its complement.  
CC Note: The sequence data for this patent did not form part  
CC of the printed specification, but was obtained in electronic  
CC format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.

XX  
SQ Sequence 7924 BP; 1973 A; 83 C; 1827 G; 4041 T; 0 other;

Query Match 16.2%; Score 36.4; DB 24; Length 7924;

Best Local Similarity 64.0%; Pred. No. 12; Mismatches 0; Indels 0; Gaps 0;

Matches 55; Conservative 0;

OY 77 CAGAAGAATATTAGTTTCACTCAGGTTTTTCAAGCTACGCTGCCCCCAAAACGAA 136

DB 3179 CAATAAATTATTACCTCTACTCATATATCAATTAACCTTCCCAATTACAAACA 3120

OY 137 ACAAAACAAAAAACAACCTTTTTTAA 162

DB 3119 ATAAAAACGAAAAAAGACCTTATTCA 3094

Search completed: June. 7, 2003, 20:37:22

Job time : 299 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 7, 2003, 17:27:51 ; Search time 898 Seconds  
(without alignments)  
7291.906 Million cell updates/sec

Title: US-09-765-231A-58  
Perfect score: 225  
Sequence: 1 tcatggttaagttgtttcagg.....attaggaattttttttttt 225

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues  
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_hgt.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pi.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_ats.\*
- 12: gb\_ey.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*
- 15: em\_ba.\*
- 16: em\_fun.\*
- 17: em\_hum.\*
- 18: em\_in.\*
- 19: em\_mu.\*
- 20: em\_om.\*
- 21: em\_or.\*
- 22: em\_ov.\*
- 23: em\_pat.\*
- 24: em\_ph.\*
- 25: em\_pi.\*
- 26: em\_ro.\*
- 27: em\_ats.\*
- 28: em\_un.\*
- 29: em\_vi.\*
- 30: em\_htg\_hum.\*
- 31: em\_htg\_inv.\*
- 32: em\_htg\_other.\*
- 33: em\_htg\_mus.\*
- 34: em\_htg\_pln.\*
- 35: em\_htg\_red.\*
- 36: em\_htg\_mam.\*
- 37: em\_htg\_vrt.\*
- 38: em\_ey.\*
- 39: em\_htgo\_hum.\*
- 40: em\_htgo\_mus.\*
- 41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	225	100.0	225	6	AX202128	Sequence
2	225	100.0	143800	2	AC011875	Homo sapi
3	225	100.0	186962	9	AP000848	Homo sapi
4	175.6	78.0	172830	2	AP001320	Homo sapi
5	128.4	57.1	75002	2	AC023384	Homo sapi
6	47.2	21.0	30726	2	AC117269	Dictyoste
7	46.2	20.5	192454	2	AC094553	Rattus no
8	45.2	20.1	9971	1	U67577	Methanococ
9	44	19.6	146353	2	AC104893	Mus muscu
10	43.8	19.5	200110	10	AC117237	Mus muscu
11	42.6	18.9	128526	2	AP005406	Oryza sat
12	42.4	18.8	199551	2	AC006281	Plasmodiu
13	42	18.7	147760	9	AC011846	Homo sapi
14	41.8	18.6	156140	9	AC067745	Homo sapi
15	41.8	18.6	215467	2	AC013420	Homo sapi
16	41.8	18.6	215734	2	AC073828	Mus muscu
17	41.6	18.5	171050	9	AC112232	Homo sapi
18	41.6	18.5	181864	2	AC027460	Homo sapi
19	41.2	18.3	235776	2	AL645950	Mus muscu
20	41	18.2	5518	6	AX323692	Sequence
21	41	18.2	169546	2	AC004157	Plasmodiu
22	40.8	18.1	62354	2	AC118028	Mus muscu
23	40.8	18.1	130355	8	AP003412	Oryza sat
24	40.6	18.0	203241	9	AC008282	Homo sapi
25	40.4	18.0	12300	10	MMU84903	Mus musculu
26	40.4	18.0	101491	10	AP003183	Mus musculu
27	40.4	18.0	107739	2	AC116979	Dictyoste
28	40.4	18.0	192107	2	AC107635	Mus muscu
29	40.4	18.0	201773	2	AC130714	Mus muscu
30	40.4	18.0	210614	2	AC013548	Mus muscu
31	40.4	18.0	252504	2	AC073815	Mus muscu
32	40.2	17.9	122032	9	AL513187	Human DNA
33	40.2	17.9	161067	2	AL607133	Homo sapi
34	40.2	17.9	208398	2	AC123245	Rattus no
35	40.2	17.9	213728	9	AL158828	Human DNA
36	40	17.8	6809	6	AX251912	Sequence
37	40	17.8	6809	6	AX344310	Sequence
38	40	17.8	6809	6	AX344766	Sequence
39	40	17.8	6809	6	AX348989	Sequence
40	40	17.8	67970	3	PFMAL1P3	Plasmodiu
41	40	17.8	68751	2	AC103861	Homo sapi
42	40	17.8	153826	2	AC083965	Homo sapi
43	40	17.8	163608	2	AC026081	Homo sapi
44	40	17.8	166379	9	AC087378	Homo sapi
45	40	17.8	176975	2	AC098753	Rattus no

ALIGNMENTS

RESULT 1  
AX202128  
LOCUS  
DEFINITION Sequence 58 from Patent WO0153531.  
ACCESSION AX202128  
VERSION AX202128.1 GI:15391919  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 225)  
AUTHORS Phippard,D., Vasanthakumr,G., Doteon,S. and Ma,X.J.  
TITLE Osteoarthritis tissue derived nucleic acids, polypeptides, vectors, and cells

225 bp DNA linear PAT 30-AUG-2001



\* 123142 143800: contig of 20659 bp in length.

```

FEATURES
  source      Location/Qualifiers
    1. .143800
       /organism="Homo sapiens"
       /db_xref="taxon:9606"
       /clone="RP11-16K5"
       /clone_lib="RPC1-11 Human Male BAC"
  misc_feature 1. .1029
               /notes="assembly_fragment"
  misc_feature 1130. .2964
               /notes="assembly_fragment"
  misc_feature 3065. .4253
               /note="assembly_fragment"
  misc_feature 4354. .5678
               /notes="assembly_fragment"
  misc_feature 5779. .6855
               /notes="assembly_fragment"
  misc_feature 6956. .8847
               /note="assembly_fragment"
  misc_feature 8948. .10945
               /notes="assembly_fragment"
  misc_feature 11046. .13001
               /notes="assembly_fragment"
  misc_feature 13102. .15430
               /note="assembly_fragment"
  misc_feature 15531. .17919
               /notes="assembly_fragment"
  misc_feature 18020. .20135
               /notes="assembly_fragment"
  misc_feature 20236. .23750
               /note="assembly_fragment"
  misc_feature 23851. .26794
               /notes="assembly_fragment"
  misc_feature 26895. .29482
               /notes="assembly_fragment"
  misc_feature 29583. .33174
               /notes="assembly_fragment"
  misc_feature 33275. .35185
               /notes="assembly_fragment"
               clone_end:SP6
               vector_side:right"
  misc_feature 35286. .39745
               /notes="assembly_fragment"
  misc_feature 39846. .44222
               /notes="assembly_fragment"
  misc_feature 44323. .48990
               /notes="assembly_fragment"
  misc_feature 49091. .57790
               /notes="assembly_fragment"
  misc_feature 57891. .66822
               /notes="assembly_fragment"
  misc_feature 66923. .76709
               /notes="assembly_fragment"
  misc_feature 76810. .92865
               /notes="assembly_fragment"
  misc_feature 92966. .106278
               /notes="assembly_fragment"
               clone_end:T7
               vector_side:right"
  misc_feature 106379. .123041
               /notes="assembly_fragment"
  misc_feature 123142. .143800
               /notes="assembly_fragment"

BASE COUNT 43074 a 27702 c 27628 g 42892 t 2504 others*
ORIGIN
Query Match      100.0%; Score 225; DB 2; Length 143800;
Best Local Similarity .100.0%; Pred. No. 5.9e-41;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGATGTTAGTCTCTTCAGGCATATAATTTGAATAAATTTAGAGCTCCATGATGCT 60
   |||||||
Db 50974 TGATGTTAGTCTCTTCAGGCATATAATTTGAATAAATTTAGAGCTCCATGATGCT 50915

FEATURES
  source      Location/Qualifiers
    1. .186962
       /organism="Homo sapiens"
       /db_xref="taxon:9606"
       /chromosomes="11"
       /maps="11q"
       /clone="RP11-775D16"
  BASE COUNT 57359 a 37204 c 36384 g 56015 t
  ORIGIN
Query Match      100.0%; Score 225; DB 9; Length 186962;
Best Local Similarity .100.0%; Pred. No. 5.7e-41;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGATGTTAGTCTCTTCAGGCATATAATTTGAATAAATTTAGAGCTCCATGATGCT 60
   |||||||
Db 83181 TGATGTTAGTCTCTTCAGGCATATAATTTGAATAAATTTAGAGCTCCATGATGCT 83122

Qy 61 ATATTGGTTTTACCTTCAGAAGAAATTTAGTTTCACTCAGGTTTTTCAAAGCTAGCTG 120
   |||||||
Db 83121 ATATTGGTTTTACCTTCAGAAGAAATTTAGTTTCACTCAGGTTTTTCAAAGCTAGCTG 83062

Qy 121 TCCCCCAAAAAACGAAACAAAAACAAAAACAAACCTTTTTTAAGAGTTGATGGTACTCAT 180
   |||||||
Db 83061 TCCCCCAAAAAACGAAACAAAAACAAAAACAAACCTTTTTTAAGAGTTGATGGTACTCAT 180

Qy 181 TTGATCTGCCTCTCTCTGCTGAATCAATTAGGAATTTTTTTTTTTT 225
   |||||||
Db 83001 TTGATCTGCCTCTCTCTGCTGAATCAATTAGGAATTTTTTTTTTTT 82957

RESULT 4

```



\* 165955 165954: gap of 100 bp  
 \* 165955 167495: contig of 1541 bp in length  
 \* 167496 167595: gap of 100 bp  
 \* 167596 168771: contig of 1176 bp in length  
 \* 168772 168871: gap of 100 bp  
 \* 168872 170322: contig of 1451 bp in length  
 \* 170323 170422: gap of 100 bp  
 \* 170423 171562: contig of 1140 bp in length  
 \* 171563 171662: gap of 100 bp  
 \* 171663 172830: contig of 1168 bp in length.

## FEATURES

source  
 1. .172830  
 /location/Qualifiers  
 /organism="Homo sapiens"  
 /db\_xrefs="taxon:9606"  
 /chromosomes="11"  
 /map="11q14"  
 /clone="RP11-79904"

misc\_feature  
 1. .15732  
 /note="assembly\_fragment"  
 misc\_feature  
 15933. .34101  
 /note="assembly\_fragment"  
 misc\_feature  
 34202. .47479  
 /note="assembly\_fragment"  
 misc\_feature  
 47580. .57756  
 /note="assembly\_fragment"  
 misc\_feature  
 57857. .69426  
 /note="assembly\_fragment"  
 misc\_feature  
 69527. .80185  
 /note="assembly\_fragment"  
 misc\_feature  
 80286. .88989  
 /note="assembly\_fragment"  
 misc\_feature  
 89090. .99362  
 /note="assembly\_fragment"  
 misc\_feature  
 9463. .103433  
 /note="assembly\_fragment"  
 misc\_feature  
 103534. .109064  
 /note="assembly\_fragment"  
 misc\_feature  
 109165. .113855  
 /note="assembly\_fragment"  
 misc\_feature  
 113956. .119140  
 /note="assembly\_fragment"  
 misc\_feature  
 119241. .124661  
 /note="assembly\_fragment"  
 misc\_feature  
 124762. .126822  
 /note="assembly\_fragment"  
 misc\_feature  
 126923. .131337  
 /note="assembly\_fragment"  
 misc\_feature  
 131438. .135078  
 /note="assembly\_fragment"  
 misc\_feature  
 135179. .138134  
 /note="assembly\_fragment"  
 misc\_feature  
 138235. .142123  
 /note="assembly\_fragment"  
 misc\_feature  
 142224. .144719  
 /note="assembly\_fragment"

Query Match 78.0%; Score 175.6; DB 2; Length 172830;  
 Best Local Similarity 88.9%; Pred. No. 6e-30;  
 Matches 201; Conservative 0; Mismatches 24; Indels 1; Gaps 1;

QY 1 TGATGGTAGTCTTTTCAGGCATATAAATTTGAATAAATATGAGCTCCATGATAGCT 60  
 Db 28909 TGATGGTAGTCTTTTCAGGCATATAAATTTGAATAAATATGAGCTCCATGATAGCT 28968  
 QY 61 ATATTGGTTTACCTTCAGAGATATTTAGTTTCACTCAGGTTTTCAGGCT-ACGCT 119  
 Db 28969 ATATTGGTTTACCTTCAGAGATATTTAGTTTCACTCAGGTTTTCAGGCTAACCCCT 29028  
 QY 120 GTCCCCCAAAACGAAACCAAAACCAAAACCAACCTTTTAAAGAGTTGATGGCTACTCA 179  
 Db 29029 GTCCCCCAAAATACGAGACCAAAACCAACCCCTTTAAAGAGTTGATGGCTAATCA 29088  
 QY 180 TTGTGATCTGCTCCTCTGTGCTGAATCAATTAGGAATTTTTTTTTTTT 225

Db 29089 TTGTGATCTGCTCCTGTGCTGATCACATTAGGATTTTCTTTT 29134

## RESULT 5

## AC023384

## LOCUS

## DEFINITION

## AC023384

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

1674 2455: contig of 782 bp in length  
2456 2555: gap of 100 bp  
2556 3328: contig of 773 bp in length  
3329 3428: gap of 100 bp  
3429 4179: contig of 751 bp in length  
4180 4279: gap of 100 bp  
4280 5022: contig of 743 bp in length  
5023 5122: gap of 100 bp  
5123 5855: contig of 743 bp in length  
5866 5965: gap of 100 bp  
5966 6703: contig of 738 bp in length  
6704 6803: gap of 100 bp  
6804 7554: contig of 751 bp in length  
7555 7654: gap of 100 bp  
7655 8408: contig of 754 bp in length  
8409 8508: gap of 100 bp  
8509 9254: contig of 746 bp in length  
9255 9354: gap of 100 bp  
9355 10118: contig of 764 bp in length  
10119 10218: gap of 100 bp  
10219 10978: contig of 760 bp in length  
10979 11078: gap of 100 bp  
11079 11850: contig of 772 bp in length  
11851 11950: gap of 100 bp  
11951 12712: contig of 762 bp in length  
12713 12812: gap of 100 bp  
12813 13553: contig of 741 bp in length  
13554 13653: gap of 100 bp  
13654 14392: contig of 739 bp in length  
14393 14492: gap of 100 bp  
14493 15244: contig of 752 bp in length  
15245 15344: gap of 100 bp  
15345 16095: contig of 751 bp in length  
16096 16195: gap of 100 bp  
16196 16927: contig of 732 bp in length  
16928 17027: gap of 100 bp  
17028 17766: contig of 739 bp in length  
17767 17866: gap of 100 bp  
17867 18615: contig of 749 bp in length  
18616 18715: gap of 100 bp  
18716 19489: contig of 774 bp in length  
19490 19589: gap of 100 bp  
19590 20344: contig of 755 bp in length  
20345 20444: gap of 100 bp  
20445 21205: contig of 761 bp in length  
21206 21305: gap of 100 bp  
21306 22089: contig of 784 bp in length  
22090 22189: gap of 100 bp  
22190 22972: contig of 783 bp in length  
22973 23072: gap of 100 bp  
23073 23826: contig of 754 bp in length  
23827 23926: gap of 100 bp  
23927 24674: contig of 748 bp in length  
24675 24774: gap of 100 bp  
24775 25507: contig of 733 bp in length  
25508 25607: gap of 100 bp  
25608 26347: contig of 740 bp in length  
26348 26447: gap of 100 bp  
26448 27202: contig of 755 bp in length  
27203 27302: gap of 100 bp  
27303 28052: contig of 750 bp in length  
28053 28152: gap of 100 bp  
28153 28909: contig of 757 bp in length  
28910 29009: gap of 100 bp  
29010 29741: contig of 732 bp in length  
29742 29841: gap of 100 bp  
29842 30624: contig of 783 bp in length  
30625 30724: gap of 100 bp  
30725 31497: contig of 773 bp in length  
31498 31597: gap of 100 bp  
31598 32373: contig of 776 bp in length  
32374 32473: gap of 100 bp  
32474 33222: contig of 749 bp in length

33223 33322: gap of 100 bp  
33323 34072: contig of 750 bp in length  
34073 34172: gap of 100 bp  
34173 34932: contig of 760 bp in length  
34933 35032: gap of 100 bp  
35033 35791: contig of 759 bp in length  
35792 35891: gap of 100 bp  
35892 36636: contig of 745 bp in length  
36637 36736: gap of 100 bp  
36737 37468: contig of 732 bp in length  
37469 37568: gap of 100 bp  
37569 38328: contig of 760 bp in length  
38329 38428: gap of 100 bp  
38429 39191: contig of 763 bp in length  
39192 39291: gap of 100 bp  
39292 40036: contig of 745 bp in length  
40037 40136: gap of 100 bp  
40137 40895: contig of 759 bp in length  
40896 40995: gap of 100 bp  
40996 41729: contig of 734 bp in length  
41730 41829: gap of 100 bp  
41830 42577: contig of 748 bp in length  
42578 42677: gap of 100 bp  
42678 43414: contig of 737 bp in length  
43415 43514: gap of 100 bp  
43515 44267: contig of 753 bp in length  
44268 44367: gap of 100 bp  
44368 45123: contig of 756 bp in length  
45124 45223: gap of 100 bp  
45224 45973: contig of 750 bp in length  
45974 46073: gap of 100 bp  
46074 46816: contig of 743 bp in length  
46817 46916: gap of 100 bp  
46917 47681: contig of 765 bp in length  
47682 47781: gap of 100 bp  
47782 48544: contig of 763 bp in length  
48545 48644: gap of 100 bp  
48645 49410: contig of 766 bp in length  
49411 49510: gap of 100 bp  
49511 50247: contig of 737 bp in length  
50248 50347: gap of 100 bp  
50348 51067: contig of 720 bp in length  
51068 51167: gap of 100 bp  
51168 51915: contig of 748 bp in length  
51916 52015: gap of 100 bp  
52016 52745: contig of 730 bp in length  
52746 52845: gap of 100 bp  
52846 53584: contig of 739 bp in length  
53585 53684: gap of 100 bp  
53685 54427: contig of 743 bp in length  
54428 54527: gap of 100 bp  
54528 55292: contig of 765 bp in length  
55293 55392: gap of 100 bp  
55393 56164: contig of 772 bp in length  
56165 56264: gap of 100 bp  
56265 56985: contig of 721 bp in length  
56986 57085: gap of 100 bp  
57086 57829: contig of 744 bp in length  
57830 57929: gap of 100 bp  
57930 58677: contig of 748 bp in length

Query Match 57.1%; Score 128.4; DB 2; Length 75002;  
Best Local Similarity 98.6%; Pred No. 2,3e-19;

Matches 140; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 84 ATATTAGTTTCACTCAGGTTTTTCAAGCTACGCTGCCCAAAACGAAACAAAC 143

DB 35932 ATATTAGTTTCACTCAGGTTTTTCAAGCTACGCTGCCCAAAACGAAAC 35990

QY 144 AAAAAACAACCTTTTAAAGAGTTGATGCTACTCATTTGATCTGCTCTGCTGAAT 203

DB 35991 AAAAAACAACCTTTTAAAGAGTTGATGCTACTCATTTGATCTGCTCTGCTGAAT 36050



**RESULT 7**  
**AC094553**

Center project name: GAWO  
Center clone name: CH230-4P5  
----- Summary Statistics

Sequencing vector: Plasmid;  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990329  
Consensus quality: 138018 bases at least Q40  
Consensus quality: 144061 bases at least Q30  
Consensus quality: 149644 bases at least Q20

-----  
\* NOTE: Estimated insert size may differ from sequence length  
(see http://www.hgsc.bcm.tmc.edu/docs/genbank\_draft\_data.html).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 56 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

\* 1 1247: contig of 1247 bp in length  
\* 1248 1347: gap of unknown length  
\* 1348 2451: contig of 1104 bp in length  
\* 2452 2551: gap of unknown length  
\* 2552 3753: contig of 1203 bp in length  
\* 3755 3854: gap of unknown length  
\* 3855 4910: contig of 1056 bp in length  
\* 4911 5010: gap of unknown length  
\* 5011 6068: contig of 1058 bp in length  
\* 6069 6169: gap of unknown length  
\* 6169 7867: contig of 1699 bp in length  
\* 7868 9573: gap of unknown length  
\* 9574 9673: contig of 1606 bp in length  
\* 9674 11193: gap of unknown length  
\* 11194 12293: gap of unknown length  
\* 12294 12334: contig of 1041 bp in length  
\* 12335 12434: gap of unknown length  
\* 12435 13442: contig of 1008 bp in length  
\* 13443 13542: gap of unknown length  
\* 13543 14578: contig of 1036 bp in length  
\* 14579 14678: gap of unknown length  
\* 14679 15894: contig of 1216 bp in length  
\* 15895 15994: gap of unknown length  
\* 15995 17042: contig of 1048 bp in length  
\* 17043 17142: gap of unknown length  
\* 17143 18366: contig of 1224 bp in length  
\* 18367 18466: gap of unknown length  
\* 18467 20240: contig of 1774 bp in length  
\* 20241 20340: gap of unknown length  
\* 20341 21680: contig of 1340 bp in length  
\* 21681 21780: gap of unknown length  
\* 21781 23022: contig of 1242 bp in length  
\* 23023 23122: gap of unknown length  
\* 23123 24457: contig of 1335 bp in length  
\* 24458 24557: gap of unknown length  
\* 24558 25926: contig of 1369 bp in length  
\* 25927 26028: gap of unknown length  
\* 26029 27481: contig of 1455 bp in length  
\* 27482 27581: gap of unknown length  
\* 27582 28600: contig of 1019 bp in length  
\* 28601 28700: gap of unknown length  
\* 28701 31205: contig of 2505 bp in length  
\* 31206 31308: gap of unknown length  
\* 31309 33785: contig of 2480 bp in length  
\* 33786 33885: gap of unknown length  
\* 33886 35594: contig of 1709 bp in length  
\* 35595 35694: gap of unknown length  
\* 35695 38190: contig of 2496 bp in length  
\* 38191 38290: gap of unknown length  
\* 38291 40684: contig of 2394 bp in length  
\* 40685 40784: gap of unknown length  
\* 40785 43033: contig of 2249 bp in length

\* 43034 43133: gap of unknown length  
\* 43134 45139: contig of 2006 bp in length  
\* 45140 45239: gap of unknown length  
\* 45240 47313: contig of 2074 bp in length  
\* 47314 47413: gap of unknown length  
\* 47414 50150: contig of 2737 bp in length  
\* 50151 50250: gap of unknown length  
\* 50251 52353: contig of 2103 bp in length  
\* 52354 52454: gap of unknown length  
\* 52455 55811: contig of 3358 bp in length  
\* 55812 55911: gap of unknown length  
\* 55912 59297: contig of 3286 bp in length  
\* 59298 62337: gap of unknown length  
\* 62338 62437: gap of unknown length  
\* 62438 66380: contig of 3943 bp in length  
\* 66381 66480: gap of unknown length  
\* 66481 70220: contig of 3740 bp in length  
\* 70221 70320: gap of unknown length  
\* 70321 73866: contig of 3546 bp in length  
\* 73867 73966: gap of unknown length  
\* 73967 76716: contig of 2749 bp in length  
\* 76717 80932: contig of 4117 bp in length  
\* 80933 81032: gap of unknown length  
\* 81033 84396: contig of 3364 bp in length  
\* 84397 84497: gap of unknown length  
\* 84498 88433: contig of 4147 bp in length  
\* 88434 88743: gap of unknown length  
\* 88744 92737: contig of 3994 bp in length  
\* 92738 92838: gap of unknown length  
\* 92839 95428: gap of unknown length  
\* 95429 101099: contig of 5671 bp in length  
\* 101100 101199: gap of unknown length  
\* 101200 106030: contig of 4831 bp in length  
\* 106031 106130: gap of unknown length  
\* 106131 111156: contig of 5026 bp in length  
\* 111157 111256: gap of unknown length  
\* 111257 115897: contig of 4641 bp in length  
\* 115898 115998: gap of unknown length  
\* 115999 120956: contig of 4959 bp in length  
\* 120957 121056: gap of unknown length  
\* 121057 127920: contig of 6864 bp in length  
\* 127921 128020: gap of unknown length  
\* 128021 134016: contig of 5996 bp in length  
\* 134017 134116: gap of unknown length  
\* 134117 143405: contig of 9289 bp in length  
\* 143406 143505: gap of unknown length  
\* 143506 151371: contig of 7766 bp in length  
\* 151372 151272: gap of unknown length  
\* 151273 162642: contig of 11271 bp in length  
\* 162643 162742: gap of unknown length  
\* 162743 173182: contig of 10440 bp in length

Query Match 20.5%; Score 46.2; DB 2; Length 192454;

Best Local Similarity 54.4%; Pred. No. 0.42; Indels 0; Gaps 0;

Matches 93; Conservative 0; Mismatches 0; Gaps 0;

QY 24 AAAATTGAAATAAATTTATGAGGCTCCATGATATGCTATATTGGTTTACCTTCAGAAGA 83  
Db 113461 AAAATTGCTTCAATTTCTTTGTATACAAAGATTGAGGCTTTATTTTCCCATAGAGAA 113520  
QY 84 ATATTTAGTTTTCACCTCAGGTTTTTCAAGCTAGCTGTCCTCCCAAAAAGAACAAAC 143  
Db 113521 TCGTTGAATTTTCACTAAGAAATTTTCAAGTACATTTAGTCTGCTGAAAAAAGAAAAG 113580  
QY 144 AAAAAAACACCTTTTAAAGAGTTGATGCTACTCATTTGATCTGCTCT 194  
Db 113581 AAAAGAGAGAAAGAAATTTCCAGTTGATTGTTACTAATAAGATTTCTCTT 113631

RESULT 8

U67577/C	9971 bp	DNA	linear	BCT 28-JAN-1998
LOCUS				
DEFINITION	Methanococcus jannaschii section 119 of 150 of the complete genome.			
ACCESSION	U67577 L77117			
VERSION	U67577.1	GI:2826400		
KEYWORDS				
SOURCE	Methanocaldococcus jannaschii.			
ORGANISM	Methanocaldococcus jannaschii			
REFERENCE	Archaea; Euryarchaeota; Methanococci; Methanococcales; Methanocaldococcaceae; Methanocaldococcus.			
AUTHORS	1 (bases 1 to 9971) Bult,C.J., White,O., Olsen,G.J., Zhou,L., Fleischmann,R.D., Sutton,G.G., Blake,J.A., FitzGerald,L.M., Clayton,R.A., Gocayne,J.D., Kerlavage,A.R., Dougherty,B.A., Tomb,J., Adams,M.D., Reich,C.I., Overbeek,R., Kirkness,E.F., Weinstock,K.G., Merrick,J.M., Glodek,A., Scott,J.D., Geoghagen,N.S., Weidman,J.F., Fuhrmann,J.L., Nguyen,D.T., Utterback,T., Kelley,J.M., Peterson,J.D., Sadow,P.W., Hanna,M.C., Cotton,M.D., Hurst,M.A., Roberts,K.M., Kaine,B., Borodovsky,M., Klenk,H.P., Fraser,C.M., Smith,H.O., Woese,C.R. and Venter,J.C.			
ITILE	Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii			
JOURNAL	Science 273 (5278), 1058-1073 (1996)			
MEDLINE	96337999			
PUBMED	8688087			
REFERENCE	2 (bases 1 to 9971) Bult,C.J., White,O., Olsen,G.J., Zhou,L., Fleischmann,R.D., Sutton,G.G., Blake,J.A., FitzGerald,L.M., Clayton,R.A., Gocayne,J.D., Kerlavage,A.R., Dougherty,B.A., Tomb,J.-F., Adams,M.D., Reich,C.I., Overbeek,R., Kirkness,E.F., Weinstock,K.G., Merrick,J.M., Glodek,A., Scott,J.D., Geoghagen,N.S., Weidman,J.F., Fuhrmann,J.L., Nguyen,D.T., Utterback,T., Kelley,J.M., Peterson,J.D., Sadow,P.W., Hanna,M.C., Cotton,M.D., Hurst,M.A., Roberts,K.M., Kaine,B., Borodovsky,M., Klenk,H.P., Fraser,C.M., Smith,H.O., Woese,C.R. and Venter,J.C.			
TITLE	Direct Submission			
JOURNAL	Submitted (27-AUG-1996) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA			
COMMENT	On Jan 30, 1998 this sequence version replaced gi:1592013.			
FEATURES	Location/Qualifiers			
source	1. .9971 /organism="Methanocaldococcus jannaschii" /db_xref="taxon:2190" 87. .554 /genes="MJ1369" 87. .554 /genes="MJ1369" /note="hypothetical protein; identified by GeneMark; putative" /codon_start=1 /transl_table=11 /product="M. jannaschii predicted coding region MJ1369" /protein_id="AAB99386.1" /db_xref="GI:1592014" /translation="MDIWTWVLFSSNIRNIEICYNIMINGFWRNAGEKQKNWR SFIRKNQIKPFDVAVQIAKGEIHAIGVETIYDDTPINDNEINLKNKTFPMRV SFSVIIISKEAVIKRPIKIQDYIGLGELEHDFNEILKAFQKFGMSIK"			
gene	584. .1570 /genes="MJ1370" 584. .1570 /genes="MJ1370" /note="hypothetical protein; identified by GeneMark; putative"			
CDS	584. .1570 /codon_start=1 /transl_table=11 /product="M. jannaschii predicted coding region MJ1370" /protein_id="AAB99387.1" /db_xref="GI:1592015" /translation="MITVGIDHGTSGITTCIKDNKKIIFKLKTELKESYLEELEK HISLEDIALTYMGDGINKILPIEKNRGVLSIEGAEKVGKGGTKVYDEIKESG LPRAVTPGHRGTECLDERFRALYSHIASPEKUSIAYAYKLFGEFNDVLSISNTV TLLIKDGFPGFDACIGAIGMLHGFIDLEMINDIAGKITANEAFSKAGAVKIAKLY KGVNTKEEIIINNYFNDCRLAVDSLLILSVSMEINSLPLLDKKNRRVVLASIGTL			
gene	3680. .4480 /genes="MJ1374" 3680. .4480 /genes="MJ1374" /note="similar to GB:AE000666 percent identity: 42.73; identified by sequence similarity; putative" /codon_start=1 /transl_table=11 /product="conserved hypothetical protein" /protein_id="AAB99381.1" /db_xref="GI:1592019" /translation="NRVEIIFLGGGGRWATITQKATGGFRIHTNLMRHVDPPGGA IURLNELKISPWRTNALFISCHPDHYTDEIIVBATQCTMKTKRGVFLGSLSVBGF GEYEVVSKYHOSKLEEVRLYPGDSAELYDTTIKATHTKGDGPPGIFGLSTIYGD I GYTSDETPQLIEDPDGVRILIANIVPKKNERIKGHLCSNDAIDLNSMKNKPELLI MNHMGKMTNPQIEAEIYSQNTGIEVIPARLGLKVELLNGRYQLIK"			
CDS	3680. .4480 /genes="MJ1374" 3680. .4480 /genes="MJ1374" /note="similar to GB:D26185 SP:P37555 PID:467446 GB:AL009126 percent identity: 23.61; identified by sequence similarity; putative"			
gene	3124. .3612 /genes="MJ1373" 3124. .3612 /genes="MJ1373" /note="hypothetical protein; identified by GeneMark; putative" /codon_start=1 /transl_table=11 /product="M. jannaschii predicted coding region MJ1373" /protein_id="AAB99388.1" /db_xref="GI:1592018" /translation="MKKTIILFLFTAILCSITLGGCISFEKTOIGNYNIKNKSCINS NVSKQTIQNVSDKIYNNNTNLNLNVEIIAYGAFGEKNRGYIYKDKNKTIIYNL BEMTAGYKIKIINIITETANKITVYKVIIPKPEFAAMVVYPIYKLSVNGTVNVECKE VR"			
CDS	3124. .3612 /genes="MJ1373" 3124. .3612 /genes="MJ1373" /note="hypothetical protein; identified by GeneMark; putative" /codon_start=1 /transl_table=11 /product="conserved hypothetical protein" /protein_id="AAB99380.1" /db_xref="GI:1592017" /translation="MGILGKIKNKLKSGIKRVIIDFYRFLDSGVULKIYEKILEBAID KNLSPKHAIVMDGNRAAEIYKORYYGHLYGAEKVRVRLWARDLGNVVTLYAFS TENFRPKEVDKMLPEKFEIADDEIHYEVVRVRAIGRLNLPKNVQAKIAYA BERTKNYKFFVIAIAYGQOEIIDAVKKAIEKVRGEIPEIDIKELIDKHLTYAN LFPNPDLIIRTSGBERISNFIWQSSYSELVFCDIYWFLFRRVDFLRAVRDQRRQR REGK"			
gene	2285. .3127 /genes="MJ1372" 2285. .3127 /genes="MJ1372" /note="similar to GB:L42023 SP:P44938 PID:1006041 PID:1221025 PID:1205167 percent identity: 47.86; identified by sequence similarity; putative" /codon_start=1 /transl_table=11 /product="conserved hypothetical protein" /protein_id="AAB99379.1" /db_xref="GI:1592016" /translation="MIIMYAIGDNKSERVLKAYEKLKEGIEVELIDNPKLVLVDKLL DGEIDKAVRGSLSKSVILYLRERIGKFRASILKNPFTNGIFLSPVGIDDISDKN ERIKRIIEFASNPLNINIKAVVLSGRLGDLGRNKVVDETIYEAEEIVEHFK GNVDIITHNGILIEEVLKCGYNIIVADVGITGNLIFCLGLICKIPGYGAVILSKNVN FIDTSRNNWERYVNAIKFLIGGDFG"			
CDS	2285. .3127 /genes="MJ1372" 2285. .3127 /genes="MJ1372" /note="similar to GP:2437836 percent identity: 39.09; identified by sequence similarity; putative" /codon_start=1 /transl_table=11 /product="methyltransferase" /protein_id="AAB99379.1" /db_xref="GI:1592016" /translation="MIIMYAIGDNKSERVLKAYEKLKEGIEVELIDNPKLVLVDKLL DGEIDKAVRGSLSKSVILYLRERIGKFRASILKNPFTNGIFLSPVGIDDISDKN ERIKRIIEFASNPLNINIKAVVLSGRLGDLGRNKVVDETIYEAEEIVEHFK GNVDIITHNGILIEEVLKCGYNIIVADVGITGNLIFCLGLICKIPGYGAVILSKNVN FIDTSRNNWERYVNAIKFLIGGDFG"			
gene	1567. .2301 /genes="MJ1371" 1567. .2301 /genes="MJ1371" /note="similar to GP:2437836 percent identity: 39.09; identified by sequence similarity; putative" /codon_start=1 /transl_table=11 /product="methyltransferase" /protein_id="AAB99379.1" /db_xref="GI:1592016" /translation="MIIMYAIGDNKSERVLKAYEKLKEGIEVELIDNPKLVLVDKLL DGEIDKAVRGSLSKSVILYLRERIGKFRASILKNPFTNGIFLSPVGIDDISDKN ERIKRIIEFASNPLNINIKAVVLSGRLGDLGRNKVVDETIYEAEEIVEHFK GNVDIITHNGILIEEVLKCGYNIIVADVGITGNLIFCLGLICKIPGYGAVILSKNVN FIDTSRNNWERYVNAIKFLIGGDFG"			
CDS	1567. .2301 /genes="MJ1371" 1567. .2301 /genes="MJ1371" /note="similar to GP:2437836 percent identity: 39.09; identified by sequence similarity; putative" /codon_start=1 /transl_table=11 /product="methyltransferase" /protein_id="AAB99379.1" /db_xref="GI:1592016" /translation="MIIMYAIGDNKSERVLKAYEKLKEGIEVELIDNPKLVLVDKLL DGEIDKAVRGSLSKSVILYLRERIGKFRASILKNPFTNGIFLSPVGIDDISDKN ERIKRIIEFASNPLNINIKAVVLSGRLGDLGRNKVVDETIYEAEEIVEHFK GNVDIITHNGILIEEVLKCGYNIIVADVGITGNLIFCLGLICKIPGYGAVILSKNVN FIDTSRNNWERYVNAIKFLIGGDFG"			
gene	1567. .2301 /genes="MJ1371" 1567. .2301 /genes="MJ1371" /note="similar to GP:2437836 percent identity: 39.09; identified by sequence similarity; putative" /codon_start=1 /transl_table=11 /product="methyltransferase" /protein_id="AAB99379.1" /db_xref="GI:1592016" /translation="MIIMYAIGDNKSERVLKAYEKLKEGIEVELIDNPKLVLVDKLL DGEIDKAVRGSLSKSVILYLRERIGKFRASILKNPFTNGIFLSPVGIDDISDKN ERIKRIIEFASNPLNINIKAVVLSGRLGDLGRNKVVDETIYEAEEIVEHFK GNVDIITHNGILIEEVLKCGYNIIVADVGITGNLIFCLGLICKIPGYGAVILSKNVN FIDTSRNNWERYVNAIKFLIGGDFG"			
CDS	1567. .2301 /genes="MJ1371" 1567. .2301 /genes="MJ1371" /note="similar to GP:2437836 percent identity: 39.09; identified by sequence similarity; putative" /codon_start=1 /transl_table=11 /product="methyltransferase" /protein_id="AAB99379.1" /db_xref="GI:1592016" /translation="MIIMYAIGDNKSERVLKAYEKLKEGIEVELIDNPKLVLVDKLL DGEIDKAVRGSLSKSVILYLRERIGKFRASILKNPFTNGIFLSPVGIDDISDKN ERIKRIIEFASNPLNINIKAVVLSGRLGDLGRNKVVDETIYEAEEIVEHFK GNVDIITHNGILIEEVLKCGYNIIVADVGITGNLIFCLGLICKIPGYGAVILSKNVN FIDTSRNNWERYVNAIKFLIGGDFG"			
gene	1567. .2301 /genes="MJ1371" 1567. .2301 /genes="MJ1371" /note="similar to GP:2437836 percent identity: 39.09; identified by sequence similarity; putative" /codon_start=1 /transl_table=11 /product="methyltransferase" /protein_id="AAB99379.1" /db_xref="GI:1592016" /translation="MIIMYAIGDNKSERVLKAYEKLKEGIEVELIDNPKLVLVDKLL DGEIDKAVRGSLSKSVILYLRERIGKFRASILKNPFTNGIFLSPVGIDDISDKN ERIKRIIEFASNPLNINIKAVVLSGRLGDLGRNKVVDETIYEAEEIVEHFK GNVDIITHNGILIEEVLKCGYNIIVADVGITGNLIFCLGLICKIPGYGAVILSKNVN FIDTSRNNWERYVNAIKFLIGGDFG"			
CDS	1567. .2301 /genes="MJ1371" 1567. .2301 /genes="MJ1371" /note="similar to GP:2437836 percent identity: 39.09; identified by sequence similarity; putative" /codon_start=1 /transl_table=11 /product="methyltransferase" /protein_id="AAB99379.1" /db_xref="GI:1592016" /translation="MIIMYAIGDNKSERVLKAYEKLKEGIEVELIDNPKLVLVDKLL DGEIDKAVRGSLSKSVILYLRERIGKFRASILKNPFTNGIFLSPVGIDDISDKN ERIKRIIEFASNPLNINIKAVVLSGRLGDLGRNKVVDETIYEAEEIVEHFK GNVDIITHNGILIEEVLKCGYNIIVADVGITGNLIFCLGLICKIPGYGAVILSKNVN FIDTSRNNWERYVNAIKFLIGGDFG"			
gene	1567. .2301 /genes="MJ1371" 1567. .2301 /genes="MJ1371" /note="similar to GP:2437836 percent identity: 39.09; identified by sequence similarity; putative" /codon_start=1 /transl_table=11 /product="methyltransferase" /protein_id="AAB99379.1" /db_xref="GI:1592016" /translation="MIIMYAIGDNKSERVLKAYEKLKEGIEVELIDNPKLVLVDKLL DGEIDKAVRGSLSKSVILYLRERIGKFRASILKNPFTNGIFLSPVGIDDISDKN ERIKRIIEFASNPLNINIKAVVLSGRLGDLGRNKVVDETIYEAEEIVEHFK GNVDIITHNGILIEEVLKCGYNIIVADVGITGNLIFCLGLICKIPGYGAVILSKNVN FIDTSRNNWERYVNAIKFLIGGDFG"			
CDS	1567. .2301 /genes="MJ1371" 1567. .2301 /genes="MJ1371" /note="similar to GP:2437836 percent identity: 39.09; identified by sequence similarity; putative" /codon_start=1 /transl_table=11 /product="methyltransferase" /protein_id="AAB99379.1" /db_xref="GI:1592016" /translation="MIIMYAIGDNKSERVLKAYEKLKEGIEVELIDNPKLVLVDKLL DGEIDKAVRGSLSKSVILYLRERIGKFRASILKNPFTNGIFLSPVGIDDISDKN ERIKRIIEFASNPLNINIKAVVLSGRLGDLGRNKVVDETIYEAEEIVEHFK GNVDIITHNGILIEEVLKCGYNIIVADVGITGNLIFCLGLICKIPGYGAVILSKNVN FIDTSRNNWERYVNAIKFLIGGDFG"			
gene	1567. .2301 /genes="MJ1371" 1567. .2301 /genes="MJ1371" /note="similar to GP:2437836 percent identity: 39.09; identified by sequence similarity; putative" /codon_start=1 /transl_table=11 /product="methyltransferase" /protein_id="AAB99379.1" /db_xref="GI:1592016" /translation="MIIMYAIGDNKSERVLKAYEKLKEGIEVELIDNPKLVLVDKLL DGEIDKAVRGSLSKSVILYLRERIGKFRASILKNPFTNGIFLSPVGIDDISDKN ERIKRIIEFASNPLNINIKAVVLSGRLGDLGRNKVVDETIYEAEEIVEHFK GNVDIITHNGILIEEVLKCGYNIIVADVGITGNLIFCLGLICKIPGYGAVILSKNVN FIDTSRNNWERYVNAIKFLIGGDFG"			
CDS	1567. .2301 /genes="MJ1371" 1567. .2301 /genes="MJ1371" /note="similar to GP:2437836 percent identity: 39.09; identified by sequence similarity; putative" /codon_start=1 /transl_table=11 /product="methyltransferase" /protein_id="AAB99379.1" /db_xref="GI:1592016" /translation="MIIMYAIGDNKSERVLKAYEKLKEGIEVELIDNPKLVLVDKLL DGEIDKAVRGSLSKSVILYLRERIGKFRASILKNPFTNGIFLSPVGIDDISDKN ERIKRIIEFASNPLNINIKAVVLSGRLGDLGRNKVVDETIYEAEEIVEHFK GNVDIITHNGILIEEVLKCGYNIIVADVGITGNLIFCLGLICKIPGYGAVILSKNVN FIDTSRNNWERYVNAIKFLIGGDFG"			
gene	1567. .2301 /genes="MJ1371" 1567. .2301 /genes="MJ1371" /note="similar to GP:2437836 percent identity: 39.09; identified by sequence similarity; putative" /codon_start=1 /transl_table=11 /product="methyltransferase" /protein_id="AAB99379.1" /db_xref="GI:1592016" /translation="MIIMYAIGDNKSERVLKAYEKLKEGIEVELIDNPKLVLVDKLL DGEIDKAVRGSLSKSVILYLRERIGKFRASILKNPFTNGIFLSPVGIDDISDKN ERIKRIIEFASNPLNINIKAVVLSGRLGDLGRNKVVDETIYEAEEIVEHFK GNVDIITHNGILIEEVLKCGYNIIVADVGITGNLIFCLGLICKIPGYGAVILSKNVN FIDTSRNNWERYVNAIKFLIGGDFG"			
CDS	1567. .2301 /genes="MJ1371" 1567. .2301 /genes="MJ1371" /note="similar to GP:2437836 percent identity: 39.09; identified by sequence similarity; putative" /codon_start=1 /transl_table=11 /product="methyltransferase" /protein_id="AAB99379.1" /db_xref="GI:1592016" /translation="MIIMYAIGDNKSERVLKAYEKLKEGIEVELIDNPKLVLVDKLL DGEIDKAVRGSLSKSVILYLRERIGKFRASILKNPFTNGIFLSPVGIDDISDKN ERIKRIIEFASNPLNINIKAVVLSGRLGDLGRNKVVDETIYEAEEIVEHFK GNVDIITHNGILIEEVLKCGYNIIVADVGITGNLIFCLGLICKIPGYGAVILSKNVN FIDTSRNNWERYVNAIKFLIGGDFG"			
gene	1567. .2301 /genes="MJ1371" 1567. .2301 /genes="MJ1371" /note="similar to GP:2437836 percent identity: 39.09; identified by sequence similarity; putative" /codon_start=1 /transl_table=11 /product="methyltransferase" /protein_id="AAB99379.1" /db_xref="GI:1592016" /translation="MIIMYAIGDNKSERVLKAYEKLKEGIEVELIDNPKLVLVDKLL DGEIDKAVRGSLSKSVILYLRERIGKFRASILKNPFTNGIFLSPVGIDDISDKN ERIKRIIEFASNPLNINIKAVVLSGRLGDLGRNKVVDETIYEAEEIVEHFK GNVDIITHNGILIEEVLKCGYNIIVADVGITGNLIFCLGLICKIPGYGAVILSKNVN FIDTSRNNWERYVNAIKFLIGGDFG"			
CDS	1567. .2301 /genes="MJ1371" 1567. .2301 /genes="MJ1371" /note="similar to GP:2437836 percent identity: 39.09; identified by sequence similarity; putative" /codon_start=1 /transl_table=11 /product="methyltransferase" /protein_id="AAB99379.1" /db_xref="GI:1592016" /translation="MIIMYAIGDNKSERVLKAYEKLKEGIEVELIDNPKLVLVDKLL DGEIDKAVRGSLSKSVILYLRERIGKFRASILKNPFTNGIFLSPVGIDDISDKN ERIKRIIEFASNPLNINIKAVVLSGRLGDLGRNKVVDETIYEAEEIVEHFK GNVDIITHNGILIEEVLKCGYNIIVADVGITGNLIFCLGLICKIPGYGAVILSKNVN FIDTSRNNWERYVNAIKFLIGGDFG"			
gene	1567. .2301 /genes="MJ1371" 1567. .2301 /genes="MJ1371" /note="similar to GP:2437836 percent identity: 39.09; identified by sequence similarity; putative" /codon_start=1 /transl_table=11 /product="methyltransferase" /protein_id="AAB99379.1" /db_xref="GI:1592016" /translation="MIIMYAIGDNKSERVLKAYEKLKEGIEVELIDNPKLVLVDKLL DGEIDKAVRGSLSKSVILYLRERIGKFRASILKNPFTNGIFLSPVGIDDISDKN ERIKRIIEFASNPLNINIKAVVLSGRLGDLGRNKVVDETIYEAEEIVEHFK GNVDIITHNGILIEEVLKCGYNIIVADVGITGNLIFCLGLICKIPGYGAVILSKNVN FIDTSRNNWERYVNAIKFLIGGDFG"			
CDS	1567. .2301 /genes="MJ1371" 1567. .2301 /genes="MJ1371" /note="similar to GP:2437836 percent identity: 39.09; identified by sequence similarity; putative" /codon_start=1 /transl_table=11 /product="methyltransferase" /protein_id="AAB99379.1" /db_xref="GI:1592016" /translation="MIIMYAIGDNKSERVLKAYEKLKEGIEVELIDNPKLVLVDKLL DGEIDKAVRGSLSKSVILYLRERIGKFRASILKNPFTNGIFLSPVGIDDISDKN ERIKRIIEFASNPLNINIKAVVLSGRLGDLGRNKVVDETIYEAEEIVEHFK GNVDIITHNGILIEEVLKCGYNIIVADVGITGNLIFCLGLICKIPGYGAVILSKNVN FIDTSRNNWERYVNAIKFLIGGDFG"			
gene	1567. .2301 /genes="MJ1371" 1567. .2301 /genes="MJ1371" /note="similar to GP:2437836 percent identity: 39.09; identified by sequence similarity; putative" /codon_start=1 /transl_table=11 /product="methyltransferase" /protein_id="AAB99379.1" /db_xref="GI:1592016" /translation="MIIMYAIGDNKSERVLKAYEKLKEGIEVELIDNPKLVLVDKLL DGEIDKAVRGSLSKSVILYLRERIGKFRASILKNPFTNGIFLSPVGIDDISDKN ERIKRIIEFASNPLNINIKAVVLSGRLGDLGRNKVVDETIYEAEEIVEHFK GNVDIITHNGILIEEVLKCGYNIIVADVGITGNLIFCLGLICKIPGYGAVILSKNVN FIDTSRNNWERYVNAIKFLIGGDFG"			
CDS	1567. .2301 /genes="MJ1371" 1567. .2301 /genes="MJ1371" /note="similar to GP:2437836 percent identity: 39.09; identified by sequence similarity; putative" /codon_start=1 /transl_table=11 /product="methyltransferase" /protein_id="AAB99379.1" /db_xref="GI:1592016" /translation="MIIMYAIGDNKSERVLKAYEKLKEGIEVELIDNPKLVLVDKLL DGEIDKAVRGSLSKSVILYLRERIGKFRASILKNPFTNGIFLSPVGIDDISDKN ERIKRIIEFASNPLNINIKAVVLSGRLGDLGRNKVVDETIYEAEEIVEHFK GNVDIITHNGILIEEVLKCGYNIIVADVGITGNLIFCLGLICKIPGYGAVILSKNVN FIDTSRNNWERYVNAIKFLIGGDFG"			
gene	1567. .2301 /genes="MJ1371" 1567. .2301 /genes="MJ1371" /note="similar to GP:2437836 percent identity: 39.09; identified by sequence similarity; putative" /codon_start=1 /transl_table=11 /product="methyltransferase" /protein_id="AAB99379.1" /db_xref="GI:1592016" /translation="MIIMYAIGDNKSERVLKAYEKLKEGIEVELIDNPKLVLVDKLL DGEIDKAVRGSLSKSVILYLRERIGKFRASILKNPFTNGIFLSPVGIDDISDKN ERIKRIIEFASNPLNINIKAVVLSGRLGDLGRNKVVDETIYEAEEIVEHFK GNVDIITHNGILIEEVLKCGYNIIVADVGITGNLIFCLGLICKIPGYGAVILSKNVN FIDTSRNNWERYVNAIKFLIGGDFG"			
CDS	1567. .2301 /genes="MJ1371" 1567. .2301 /genes="MJ1371" /note="similar to GP:2437836 percent identity: 39.09; identified by sequence similarity; putative" /codon_start=1 /transl_table=11 /product="methyltransferase" /protein_id="AAB99379.1" /db_xref="GI:1592016" /translation="MIIMYAIGDNKSERVLKAYEKLKEGIEVELIDNPKLVLVDKLL DGEIDKAVRGSLSKSVILYLRERIGKFRASILKNPFTNGIFLSPVGIDDISDKN ERIKRIIEFASNPLNINIKAVVLSGRLGDLGRNKVVDETIYEAEEIVEHFK GNVDIITHNGILIEEVLKCGYNIIVADVGITGNLIFCLGLICKIPGYGAVILSKNVN FIDTSRNNWERYVNAIKFLIGGDFG"			
gene	1567. .2301 /genes="MJ1371" 1567. .2301 /genes="MJ1371" /note="similar to GP:2437836 percent identity: 39.09; identified by sequence similarity; putative" /codon_start=1 /transl_table=11 /product="methyltransferase" /protein_id="AAB99379.1" /db_xref="GI:1592016" /translation="MIIMYAIGDNKSERVLKAYEKLKEGIEVELIDNPKLVLVDKLL DGEIDKAVRGSLSKSVILYLRERIGKFRASILKNPFTNGIFLSPVGIDDISDKN ERIKRIIEFASNPLNINIKAVVLSGRLGDLGRNKVVDETIYEAEEIVEHFK GNVDIITHNGILIEEVLKCGYNIIVADVGITGNLIFCLGLICKIPGYGAVILSKNVN FIDTSRNNWERYVNAIKFLIGGDFG"			
CDS	1567. .2301 /genes="MJ1371" 1567. .2301 /genes="MJ1371" /note="similar to GP:2437836 percent identity: 39.09; identified by sequence similarity; putative" /codon_start=1 /transl_table=11 /product="methyltransferase" /protein_id="AAB99379.1" /db_xref="GI:1592016" /translation="MIIMYAIGDNKSERVLKAYEKLKEGIEVELIDNPKLVLVDKLL DGEIDKAVRGSLSKSVILYLRERIGKFRASILKNPFTNGIFLSPVGIDDISDKN ERIKRIIEFASNPLNINIKAVVLSGRLGDLGRNKVVDETIYEAEEIVEHFK GNVDIITHNGILIEEVLKCGYNIIVADVGITGNLIFCLGLICKIPGYGAVILSKNVN FIDTSRNNWERYVNAIKFLIGGDFG"			
gene	1567. .2301 /genes="MJ1371" 1567. .2301 /genes="MJ1371" /note="similar to GP:2437836 percent identity: 39.09; identified by sequence similarity; putative" /codon_start=1 /transl_table=11 /product="methyltransferase" /protein_id="AAB99379.1" /db_xref="GI:1592016" /translation="MIIMYAIGDNKSERVLKAYEKLKEGIEVELIDNPKLVLVDKLL DGEIDKAVRGSLSKSVILYLRERIGKFRASILKNPFTNGIFLSPVGIDDISDKN ERIKRIIEFASNPLNINIKAVVLSGRLGDLGRNKVVDETIYEAEEIVEHFK GNVDIITHNGILIEEVLKCGYNIIVADVGITGNLIFCLGLICKIPGYGAVILSKNVN FIDTSRNNWERYVNAIKFLIGGDFG"			
CDS	1567. .2301 /genes="MJ1371" 1567. .2301 /genes="MJ1371" /note="similar to GP:2437836 percent identity: 39.09; identified by sequence similarity; putative" /codon_start=1 /transl_table=11 /product="methyltransferase" /protein_id="AAB99379.1" /db_xref="GI:1592016" /translation="MIIMYAIGDNKSERVLKAYEKLKEGIEVELIDNPKLVLVDKLL DGEIDKAVRGSLSKSVILYLRERIGKFRASILKNPFTNGIFLSPVGIDDISDKN ERIKRIIEFASNPLNINIKAVVLSGRLGDLGRNKVVDETIYEAEEIVEHFK GNVDIITHNGILIEEVLKCGYNIIVADVGITGNLIFCLGLICKIPGYGAVILSKNVN FIDTSRNNWERYVNAIKFLIGGDFG"			
gene	1567. .2301 /genes="MJ1371" 1567. .2301 /genes="MJ1371" /note="similar to GP:2437836 percent identity: 39.09; identified by sequence similarity; putative" /codon_start=1 /transl_table=11 /product="methyltransferase" /protein_id="AAB99379.1" /db_xref="GI:1592016" /translation="MIIMYAIGDNKSERVLKAYEKLKEGIEVELIDNPKLVLVDKLL DGEIDKAVRGSLSKSVILYLRERIGKFRASILKNPFTNGIFLSPVGIDDISDKN ERIKRIIEFASNPLNINIKAVVLSGRLGDLGRNKVVDETIYEAEEIVEHFK GNVDIITHNGILIEEVLKCGYNIIVADVGITGNLIFCLGLICKIPGYGAVILSKNVN FIDTSRNNWERYVNAIKFLIGGDFG"			
CDS	1567. .2301 /genes="MJ1371" 1567. .2301 /genes="MJ1371" /note="similar to GP:2437836 percent identity: 39.09; identified by sequence similarity; putative" /codon_start=1 /transl_table=11 /product="methyltransferase" /protein_id="AAB99379.1" /db_xref="GI:1592016" /translation="MIIMYAIGDNKSERVLKAYEKLKEGIEVELIDNPKLVLVDKLL DGEIDKAVRGSLSKSVILYLRERIGKFRASILKNPFTNGIFLSPVGIDDISDKN ERIKRIIEFASNPLNINIKAVVLSGRLGDLGRNKVVDETIYEAEEIVEHFK GNVDIITHNGILIEEVLKCGYNIIVADVGITGNLIFCLGLICKIPGYGAVILSKNVN FIDTSRNNWERYVNAIKFLIGGDFG"			
gene	1567. .2301 /genes="MJ1371" 1567. .2301 /genes="MJ1371" /note="similar to GP:2437836 percent identity: 39.09; identified by sequence similarity; putative" /codon_start=1 /transl_table=11 /product="methyltransferase" /protein_id="AAB99379.1" /db_xref="GI:1592016" /translation="MIIMYAIGDNKSERVLKAYEKLKEGIEVELIDNPKLVLVDKLL DGEIDKAVRGSLSKSVILYLRERIGKFRASILKNPFTNGIFLSPVGIDDISDKN ERIKRIIEFASNPLNINIKAVVLSGRLGDLGRNKVVDETIYEAEEIVEHFK GNVDIITHNGILIEEVLKCGYNIIVADVGITGNLIFCLGLICKIPGYGAVILSKNVN FIDTSRNNWERYVNAIKFLIGGDFG"			
CDS	1567. .2301 /genes="MJ1371" 1567. .2301 /genes="MJ1371" /note="similar to GP:2437836 percent identity: 39.09; identified by sequence similarity; putative" /codon_start=1 /transl_table=11 /product="methyltransferase" /protein_id="AAB99379.1" /db_xref="GI:1592016" /translation="MIIMYAIGDNKSERVLKAYEKLKEGIEVELIDNPKLVLVDKLL DGEIDKAVRGSLSKSVILYLRERIGKFRASILKNPFTNGIFLSPVGIDDISDKN ERIKRIIEFASNPLNINIKAVVLSGRLGDLGRNKVVDETIYEAEEIVEHFK GNVDIITHNGILIEEVLKCGYNIIVADVGITGNLIFCLGLICKIPGYGAVILSKNVN FIDTSRNNWERYVNAIKFLIGGDFG"			
gene	1567. .2301 /genes="MJ1371" 1567. .2301 /genes="MJ1371" /note="similar to GP:2437836 percent identity: 39.09; identified by sequence similarity; putative" /codon_start=1 /transl_table=11 /product="methyltransferase" /protein_id="AAB99379.1" /db_xref="GI:1592016" /translation="MIIMYAIGDNKSERVLKAYEKLKEGIEVELIDNPKLVLVDKLL DGEIDKAVRGSLSKSVILYLRERIGKFRASILKNPFTNGIFLSPVGIDDISDKN ERIKRIIEFASNPLNINIKAVVLSGRLGDLGRNKVVDETIYEAEEIVEHFK GNVDIITHNGILIEEVLKCGYNIIVADVGITGNLIFCLGLICKIPGYGAVILSKNVN FIDTSRNNWERYVNAIKFLIGGDFG"			
CDS	1567. .2301 /genes="MJ1371" 1567. .2301 /			



All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)

----- Project Information

Center project name: L19047

Center clone name: 288 O 15

----- Summary Statistics

Sequencing vector: Plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 142944 bases at least Q40

Consensus quality: 14423 bases at least Q30

Consensus quality: 144956 bases at least Q20

Insert size: 152000; agarose-fp

Insert size: 145353; sum-of-contigs

Quality coverage: 6.1 in Q20 bases; agarose-fp

Quality coverage: 6.4 in Q20 bases; sum-of-contigs

\* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 11 contigs. Gaps between the contigs  
 \* are represented as runs of N. The order of the pieces  
 \* is believed to be correct as given, however the sizes  
 \* of the gaps between them are based on estimates that have  
 \* provided by the submitter.

\* This sequence will be replaced  
 \* by the finished sequence as soon as it is available and  
 \* the accession number will be preserved.

\* 1 872: contig of 872 bp in length

\* 873 972: gap of 100 bp

\* 973 4027: contig of 3055 bp in length

\* 4028 4127: gap of 100 bp

\* 4128 7985: contig of 3838 bp in length

\* 7966 8065: gap of 100 bp

\* 8066 15054: contig of 6989 bp in length

\* 15055 15154: gap of 100 bp

\* 15155 22231: contig of 7077 bp in length

\* 22232 22331: gap of 100 bp

\* 22332 33166: contig of 10835 bp in length

\* 33167 33266: gap of 100 bp

\* 33267 44383: contig of 11117 bp in length

\* 44384 44483: gap of 100 bp

\* 44484 56097: contig of 11614 bp in length

\* 56098 56197: gap of 100 bp

\* 56198 83111: contig of 26914 bp in length

\* 83112 83211: gap of 100 bp

\* 83212 113014: contig of 29803 bp in length

\* 113015 113114: gap of 100 bp

\* 113115 146353: contig of 33239 bp in length.

#### FEATURES

##### source

1. .146353

/organism="Mus musculus"

/db xref="taxon:10090"

/clone="RP23-288015"

/clone\_lib="RPCI-23 Female Mouse BAC"

1. .872

/notes="assembly\_fragment"

973. .4027

/notes="assembly\_fragment"

4128. .7965

/notes="assembly\_fragment"

8066. .15054

/notes="assembly\_fragment"

15155. .22231

/notes="assembly\_fragment"

22332. .33166

/notes="assembly\_fragment"

33267. .44383

/notes="assembly\_fragment"

misc\_feature 44484. .56097

/note="assembly\_fragment"

misc\_feature 56198. .83111

/note="assembly\_fragment"

misc\_feature 83212. .113014

/note="assembly\_fragment"

misc\_feature 113115. .146353

/note="assembly\_fragment"

BASE COUNT 46768 a 27185 c 26808 g 44590 t 1002 others

ORIGIN

Query Match 19.6%; Score 44; DB 2; Length 146353;

Best Local Similarity 51.5%; Pred. No. 1.4;

Matches 101; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

QY 23 TAAATTTTCAAAATAAATTATGAGGTCCTCATGATATCTATATTGGTTTACCTTCAGAAG 82

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 89100 TAAATAATGAATAAATAAATAAATATACACATTTTCCAAATAAAGACACATCTTTATAAA 89041

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

QY 83 ATATTTAGTTTCACTCAGGTTTTTCAAGGTACGCTGTCTCCCAAAACGAAACAAAA 142

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 89040 AAAATTTAGGAACAGTCAGGTTTATACAGAGAAACCTGTCTCAAAAAACAAACAAAC 88981

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

QY 143 CAAAAACAACTTTTAAAGAGTTGATGCTACTCATTTGATCTCGCTCTCTGCTGAA 202

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 88980 AACAAACAAATCAATTTAATTATGCAATAGCTCCTATTCTTTATCCTATCATGCT 88921

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

QY 203 TCAATTTAGGAATTTT 218

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 88920 AAATAAATAATTTT 88905

RESULT 10

AC117237/c

LOCUS

AC117237

DEFINITION

AC117237

ACCESSION

AC117237.2

VERSION

HTG.

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AC117237 200110 bp DNA linear ROD 16-JUL-2002  
 Mus musculus clone RP23-431G18, complete sequence.

AC117237 GI:21206431

HTG.

house mouse.

Mus musculus

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 200110)

McPherson,J.D. and Waterston,R.H.

The sequence of Mus musculus clone

Unpublished

2 (bases 1 to 200110)

McPherson,J.D. and Waterston,R.H.

Direct Submission

Submitted (08-APR-2002) Genome Sequencing Center, 4444 Forest Park

Parkway, St. Louis, MO 63108, USA

3 (bases 1 to 200110)

McPherson,J.D. and Waterston,R.H.

Direct Submission

Submitted (25-MAY-2002) Genome Sequencing Center, 4444 Forest Park

Parkway, St. Louis, MO 63108, USA

4 (bases 1 to 200110)

McPherson,J.D. and Waterston,R.H.

Direct Submission

Submitted (16-JUL-2002) Genome Sequencing Center, 4444 Forest Park

Parkway, St. Louis, MO 63108, USA

On May 25, 2002 this sequence version replaced gi:20069804.

----- Genome Center -----

Center: Washington University Genome Sequencing Center

Center code: WUGSC

Web site: <http://genome.wustl.edu/gsc/index.shtml>

Contact: [submissions@wustl.edu](mailto:submissions@wustl.edu)

----- Project Information -----

Center project name: M\_BA0431G18

----- Location/Qualifiers -----

FEATURES

```

source      1. .200110
            /organism="Mus musculus"
            /db xref="taxon:10090"
            /clone="RP23-431G18"
BASE COUNT  60329 a 37978 c 38823 g 62980 t
ORIGIN
Query Match      19.5%; Score 43.8; DB 10; Length 200110;
Best Local Similarity 55.6%; Pred. No. 1.4;
Matches 84; Conservative 0; Mismatches 67; Indels 0; Gaps 0;
QY 75 TTCAAGAAATATTAGTTTCATCAGGTTTTCACAAAGCTACGCTGCCCAAAAACG 134
    |||||
Db 105570 TACAGAGTGAAGTCCAGGAGCAGCGGTACACAGAGAAACCTGCTCCAAAACAAA 105511
QY 135 ARAACAACAAAAAACAACCTTTTAAAGATTGATGGCTACTCAATTTGATCTGCCTCT 194
    |||||
Db 105510 AAAAAAACAACCAACCAAGAAATATTTTGAATGGGAACAGTTTATATCTTAATTTA 105451
QY 195 CTGCTGAATCAATTAGGAATTTTTTTTTTTT 225
    |||||
Db 105450 TTATTTTACATTAGGTGTTATGTGTTT 105420

RESULT 11
AP005406/c      128526 bp      DNA      linear      HTG 13-JUN-2002
LOCUS
DEFINITION      Oryza sativa (japonica cultivar-group) chromosome 8 clone B1147B12,
*** SEQUENCING IN PROGRESS ***, in ordered pieces.
ACCESSION      AP005406
VERSION         AP005406.1 GI:21396516
KEYWORDS        HTG; HTGS PHASE2.
SOURCE          Oryza sativa (japonica cultivar-group) (cultivar:Nipponbare) DNA,
clone:B1147B12.
ORGANISM        Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
REFERENCE       1
AUTHORS         Sasaki, T., Matsumoto, T. and Katayose, Y.
TITLE           Oryza sativa nippobare (GA3) genomic DNA, chromosome 8, BAC
clone:B1147B12
JOURNAL         Published Only in Database (2002)
REFERENCE       2 (bases 1 to 128526)
AUTHORS         Sasaki, T., Matsumoto, T. and Katayose, Y.
TITLE           Direct Submission
JOURNAL         Submitted (12-JUN-2002) Takuji Sasaki, National Institute of
Agrobiological Sciences, Rice Genome Research Program; Kannondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
Tel:81-298-38-7441, Fax:81-298-38-7468)
COMMENT         NOTE: It currently consists of 1 contigs. Gaps between the contigs
are represented as runs of N. The order of the pieces is believed
to be correct as given, however the sizes of the gaps between them
are based on estimates that have provided by the submitter. This
sequence will be replaced by the finished sequence as soon as it is
available and the accession number will be preserved.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
FEATURES       source
               1. .128526
                /organism="Oryza sativa (japonica cultivar-group)"
                /cultivar="Nipponbare"
                /db xref="taxon:39947"
                /chromosome="8"
                /clone="B1147B12"
BASE COUNT     37912 a 25748 c 26754 g 37999 t 113 others
ORIGIN
Query Match      18.9%; Score 42.6; DB 2; Length 128526;
Best Local Similarity 51.9%; Pred. No. 2.9;

```

```

Matches 96; Conservative 0; Mismatches 89; Indels 0; Gaps 0;
QY 30 TCAATTAATATGAGGCTCCATGATATGCTATATTTGTTTACCTTCAGAGAATATTT 89
    |||||
Db 35653 TGAATAGCTGTGATATCCAGGAAATCGTCATCATTTTCAAGACATTTGGATGATAG 35594
QY 90 AGTTTCATCAGGTTTTTCAAGCTACGCTGCTCCCAAAAACGAAACCAACAAAA 149
    |||||
Db 35593 AATCCAGAGAGCTCAAAGTTTGCATTGCTGACCCCAAAAAGAAACCAACAAA 35534
QY 150 ACAACCTTTTAAAGATTGATGCTACTCATTTGATCTGCCTCTGCTGAATTA 209
    |||||
Db 35533 GCATACATGAAAAATGTCAATGATAAGTTAATAAGATGGTCAACGCGATAATCCAACCA 35474
QY 210 GGAAT 214
    |||||
Db 35473 CAATT 35469

RESULT 12
AC006281/c      199551 bp      DNA      linear      HTG 12-AUG-2000
LOCUS
DEFINITION      Plasmodium falciparum chromosome 12 clone 3D7, *** SEQUENCING IN
PROGRESS ***, 2 unordered pieces.
ACCESSION      AC006281
VERSION         AC006281.8 GI:97977738
KEYWORDS        HTG; HTGS PHASE1.
SOURCE          Plasmodium falciparum.
ORGANISM        Plasmodium falciparum
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE       1 (bases 1 to 199551)
AUTHORS         Hyman, R.W., Fung, E.L., Qin, F., Rowley, D., Mao, J., Tamaki, T.,
Kurdi, O.B., Conway, A.B. and Davis, R.W.
TITLE           Plasmodium falciparum 3D7 chromosome 12
JOURNAL         Unpublished
REFERENCE       2 (bases 1 to 199551)
AUTHORS         Hyman, R.W., Fung, E.L., Qin, F., Tamaki, T., Kurdi, O.B. and Davis, R.W.
TITLE           Direct Submission
JOURNAL         Submitted (05-JAN-1999) Stanford DNA Sequencing and Technology
Center, Stanford University, 855 California Avenue, Palo-Alto, CA
94304, USA
COMMENT         On Aug 12, 2000 this sequence version replaced gi:8810457.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 41526: contig of 41526 bp in length
* 41527 41726: gap of unknown length
* 41727 199551: contig of 157825 bp in length.
FEATURES       source
               1. .199551
                /organism="Plasmodium falciparum"
                /db xref="taxon:5833"
                /chromosome="12"
                /clone="PFYACB8-95"
                /clone="3D7"
BASE COUNT     83197 a 17543 c 19759 g 78852 t 200 others
ORIGIN
Query Match      18.8%; Score 42.4; DB 2; Length 199551;
Best Local Similarity 50.5%; Pred. No. 2.9;
Matches 103; Conservative 0; Mismatches 101; Indels 0; Gaps 0;
QY 22 ATAAATTTGAAATAAATATGAGGCTCCATGATGCTATATTTGTTTACCTTCAGAA 81
    |||||
Db 22946 AAAAAAATAAATAAACAACATATTTTATATATATATATATATATATGTAAG 22887
QY 82 GAATATTTAGTTTCTACTAGGTTTTTCAAGCTACGCTGCTCCCAAAAACGAAACAA 141
    |||||

```

```

Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
Chazaro,B., Chospel,Y., Collangelo,M., Collins,S., Collymore,A.,
Cook,A., Cooke,P., DeArillano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., Fitzgerald,M., FitzHugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Hegos,B., Horton,L., Hulme,W., Iliev,I.,
Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,
Lamarez,R., Landers,T., Lehoczyk,J., Levine,R., Lindblad-Toh,K.,
Liu,G., Maclean,C., Macdonald,P., Major,J., Marquis,N.,
Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrim,J.,
Meneus,L., Mihova,T., Menga,V., Murphy,T., Naylor,J., Nguyen,C.,
Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
Pollara,V., Raymond,C., Ratta,R., Rieback,M., Riley,R., Rise,C.,
Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S.,
Schuchback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S.,
Theodore,J., Toham,K., Travers,M., Travis,N., Triglio,J., Tesfaye,S.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zembek,U., Zimmer,A. and Zody,M.
Direct Submission
Submitted (11-JUN-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jun 11, 2002 this sequence version repeated gi:14994158.
All repeats were identified using RepeatMasker:
[Smith, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
-----
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
-----
Project Information
Center project name: L2618
Center clone name: 346_A_8
-----

Only the last 147.8 kb of this clone are being submitted.
The remainder overlaps accession number AC105129 [WICGR project
L22971].

Location/Qualifiers
1. 14760
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="15"
/map="15"
/clone="RP11-346A8"
/clone_lib="RPC1-11 Human Male BAC"
572..672
/rpt_family="MIR"
697..1138
/rpt_family="L1MB6"
1143..1340
/rpt_family="AluJo"
1357..1665
/rpt_family="AluSg"
1666..1699
/rpt_family="L2"
1771..2026
/rpt_family="AluJo"
2049..2069
/rpt_family="(TC)n"
2072..2096
/rpt_family="(CA)n"
2134..2185
/rpt_family="(TG)n"
complement(2458..2533)
/rpt_family="MIR3"
2943..3081
/rpt_family="MIR3"
complement(3473..3526)
/rpt_family="L2"
3584..3758

```

project

repeat\_region /rpt family="MIR"  
complement(3767..4021)  
repeat\_region /rpt family="AluSx"  
complement(4075..4115)  
repeat\_region /rpt family="AluSx"  
complement(4737..4774)  
repeat\_region /rpt family="MIR"  
5062..5168  
repeat\_region /rpt family="MIR"  
complement(5262..5563)  
repeat\_region /rpt family="AluY"  
complement(6726..7022)  
repeat\_region /rpt family="AluSx"  
7049..7166  
repeat\_region /rpt family="L2"  
7215..7332  
repeat\_region /rpt family="MIR"  
complement(7387..7530)  
repeat\_region /rpt family="L3"  
8050..8168  
repeat\_region /rpt family="MIR3"  
complement(8169..8480)  
repeat\_region /rpt family="AluJo"  
8481..8537  
repeat\_region /rpt family="MIR3"  
complement(9062..9373)  
repeat\_region /rpt family="AluSx"  
9383..9601  
repeat\_region /rpt family="L2"  
9755..10149  
repeat\_region /rpt family="L2"  
11647..11688  
repeat\_region /rpt family="(CA)n"  
11360..11991  
repeat\_region /rpt family="GC-rich"  
12008..12028  
repeat\_region /rpt family="(CCG)n"  
12084..12161  
repeat\_region /rpt family="G-rich"  
complement(12722..12877)  
repeat\_region /rpt family="MIR3"  
13023..13086  
repeat\_region /rpt family="L3"  
13212..13282  
repeat\_region /rpt family="L2"  
13283..13537  
repeat\_region /rpt family="L2"  
13538..13667  
repeat\_region /rpt family="FLAM\_C"  
13668..14227  
repeat\_region /rpt family="L2"  
complement(14228..14513)  
repeat\_region /rpt family="AluSx"  
14514..14525  
repeat\_region /rpt family="L2"  
complement(14526..14835)  
repeat\_region /rpt family="AluSx"  
14836..14999  
repeat\_region /rpt family="L2"  
complement(15356..15637)  
repeat\_region /rpt family="L2"  
15859..15991  
repeat\_region /rpt family="MIR3"  
16064..16363  
repeat\_region /rpt family="AluSx"  
16597..16909  
repeat\_region /rpt family="AluJb"

Query Match 18.7%; Score 42; DB 9; Length 147760;  
Best Local Similarity 52.9%; Pred.No.3.8;  
Matches 90; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 17 CAGGCATAAAATTTGAAATAAATATGAGGCTCCATGATATGCTATATGTTGGTTTACCTT 76

Db 107925 CACTAATAAAAAATCCATTTGGGTACTATGCTCCACTCCACAGTATATAATTTATTCACCTC 107984  
QY 77 CAGAGAATATTTAGTTTTCACTCAGGTTTTTCAAGAGCTACGCTGTCCTCCCAAAAAACGAA 136  
Db 107985 TGAGGTTAAATGAGGCTTCCCTCAGGCTTTTAAAGTTGAGCTCCGTTAAAAAATAA 108044  
QY 137 ACAAAACAAAAAACAACCTTTTAAAGAGTTGATGGCTACTCATTTGATC 186  
Db 108045 AAAAAAAGGCAAGTCATTTTCTTGGTCTCTTCTTATTAATTTTCTC 108094  
RESULT 14  
AC067745/c 156140 bp DNA linear PRI 10-MAY-2002  
LOCUS Homo sapiens chromosome 10 clone RP11-417C21, complete sequence.  
DEFINITION AC067745  
AC067745  
VERSION AC067745.7 GI:20522196  
KEYWORDS HTG.  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 156140)  
AUTHORS Smith,D.R.  
TITLE Genome Therapeutics Corporation Sequencing Center: Human Genome  
Sequence Data  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 156140)  
AUTHORS Smith,D.R.  
TITLE Direct Submission  
JOURNAL Submitted (27-APR-2000) Genome Therapeutics Corporation, 100 Beaver  
Street, Waltham, MA 02453, USA  
REFERENCE 3 (bases 1 to 156140)  
AUTHORS Smith,D.R.  
TITLE Direct Submission  
JOURNAL Submitted (10-MAY-2002) Genome Therapeutics Corporation, 100 Beaver  
Street, Waltham, MA 02453, USA  
COMMENT On May 10, 2002 this sequence version replaced gi:13811871.  
FEATURES  
source  
1..156140  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="10"  
/clone="RP11-417C21"  
/clone\_lib="RPCI-11"  
BASE COUNT 44203 a 38631 c 34570 g 38736 t  
ORIGIN

Query Match 18.6%; Score 41.8; DB 9; Length 156140;  
Best Local Similarity 50.2%; Pred.No.4.2;  
Matches 103; Conservative 0; Mismatches 102; Indels 0; Gaps 0;  
QY 21 CATAAAATTTGAAATAAATATGAGGCTCCATGATATGCTATATGTTGGTTTACCTTCAGA 80  
Db 139039 CGTTTAATTCGTTTATATCTAATACTACATGATACATATGTCGTGTTTCTGCTGT 138980  
QY 81 AGAATATTTAGTTTCACCTCAGGTTTTTCAAGCTACGCTGTCCTCCCAAAAAACGAAACAA 140  
Db 138979 TTTATTTTGGGTTTCTTCTGCTCTCTTTCCCTCTCTTATTAATTAAGAAATTTGT 138920  
QY 141 ACAAAAAAACAACCTTTTAAAGAGTTGATGGCTACTCATTTGATCTGCTCTCTGCTG 200  
Db 138919 TCCATATTTCTTCTTTTAAATTTGTTGGTTTCATAGCTAGATCATATTCGTTC 138860  
QY 201 AATCAATTAGGAATTTTTTTTTTTT 225  
Db 138859 TCCTAGTGATTACTCTTTTTTTTTT 138835

RESULT 15  
AC013420  
LOCUS

AC013420 215467 bp DNA linear HTG 27-AUG-2000



DEFINITION Homo sapiens clone RP11-619F23, WORKING DRAFT SEQUENCE, 19  
 AC013420  
 AC013420.13 GI:9929508  
 HTG: HTGS\_PHASE1; HTGS\_DRAFT.  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens

REFERENCE  
 AUTHORS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 215467)  
 Muzny, D.M., Adams, C., Baile, M., Barbara, J., Blankenburg, K.,  
 Budota, B., Bouck, J., Bowie, S., Brooks, A., Buhay, C., Bunac, C.,  
 Burkett, C., Burrows, J., Carter, M., Chacko, J., Chen, Z., Cox, C.,  
 David, R., Delgado, O., Deshazo, D., Ding, Y., Domah-Rashid, N.,  
 Dugan-Rocha, S., Durbin, K.J., Fernandez, C., Ferraguto, D.,  
 Forcum-Taney, J., Frantz, P., Ganesh, R., Gorrell, J.H., Gorrell, L.L.,  
 Guevara, W., Harris, K., Hernandez, J., Hodgson, A., Hogues, M.,  
 Holloway, C., Hosak, H., Jackson, L.E., Jackson, L., Jia, Y., Jones, M.,  
 Kelly, S., Kondejewski, N., Kong, Y., Kovar, C., Leal, B., Li, Z.,  
 Lichtarge, O., Liu, J., Liu, W., Logan, O., Lozado, R.J., Lu, J.,  
 Lucier, R., Martin, R., Martinez, C., McLeod, M.P., Mei, G., Morgan, M.,  
 Morris, S., Nash, S., Nelson, A., Nguyen, R., Nguyen, N., Nguyen, S.,  
 Oswal, G., Parish, B., Paxton, S., Payton, B., Perez, L., Pu, L.L.,  
 Quiles, M., Reiter, D., Rives, M., Samuel, S., Say, J., Scherer, S.,  
 Shah, E., Shen, H., Simon, M., Sparks, A., Stamps, A., Sucgang, R.,  
 Tabor, P., Taylor, T., Vasquez, L., Vinson, R., Vo, Q., Wahbah, M.,  
 Watlington, S., Weinstock, G., Weinstock, I.R., Williamson, M.,  
 Worley, K., Wren, J., Wrensford, G., Yu, W., Zhou, X., Nelson, D. and  
 Gibbs, R.

TITLE Direct Submission  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 215467)  
 AUTHORS Worley, K.C.

TITLE Direct Submission  
 JOURNAL Submitted (10-NOV-1999) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 On Aug 26, 2000 this sequence version replaced gi:9719539.  
 ----- Genome Center  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: <http://www.hgsc.bcm.tmc.edu/>  
 Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
 ----- Project Information  
 Center project name: HMMP  
 Center clone name: RP11-619F23  
 ----- Summary Statistics  
 Sequencing vector: M13; L08821  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.990329  
 Consensus quality: 175806 bases at least Q40  
 Consensus quality: 195338 bases at least Q30  
 Consensus quality: 202378 bases at least Q20  
 Estimated insert size: 207191; sum-of-contigs estimation  
 Estimated insert size: 249405; agarose-fp estimation  
 Quality coverage: 2.7x in Q20 bases; agarose-fp estimation  
 Quality coverage: 3.3x in Q20 bases; sum-of-contigs estimation  
 -----  
 \* NOTE: Estimated insert size may differ from sequence length  
 (see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).  
 \* NOTE: This is a 'working draft' sequence. It currently  
 consists of 19 contigs. The true order of the pieces  
 is not known and their order in this sequence record is  
 arbitrary. Gaps between the contigs are represented as  
 runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 as soon as it is available and the accession number will  
 be preserved.  
 \* 1 30720: contig of 30720 bp in length  
 \* 30721 30820: gap of unknown length  
 \* 30821 57766: contig of 26946 bp in length  
 \* 57767 57866: gap of unknown length  
 \* 57867 79135: contig of 21269 bp in length

79136 79235: gap of unknown length  
 79236 93871: contig of 14636 bp in length  
 93872 93971: gap of unknown length  
 93972 109687: contig of 15716 bp in length  
 109688 109787: gap of unknown length  
 109788 124632: contig of 14845 bp in length  
 124633 124732: gap of unknown length  
 124733 139346: contig of 14614 bp in length  
 139347 139446: gap of unknown length  
 139447 150488: contig of 11042 bp in length  
 150489 150588: gap of unknown length  
 150589 162913: contig of 12325 bp in length  
 162914 163013: gap of unknown length  
 163014 171860: contig of 8647 bp in length  
 171861 171860: gap of unknown length  
 171861 178757: contig of 6996 bp in length  
 178758 178956: gap of unknown length  
 178957 186909: contig of 8053 bp in length  
 186910 187009: gap of unknown length  
 187010 194538: contig of 7629 bp in length  
 194539 194738: gap of unknown length  
 194739 199508: contig of 4770 bp in length  
 199509 199608: gap of unknown length  
 199609 204158: contig of 4550 bp in length  
 204159 204258: gap of unknown length  
 204259 207133: contig of 2875 bp in length  
 207134 207233: gap of unknown length  
 207234 210635: contig of 3302 bp in length  
 210636 214187: contig of 3552 bp in length  
 214188 214287: gap of unknown length  
 214288 215467: contig of 1180 bp in length.

FEATURES  
 Location/Qualifiers  
 1..215467  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="RP11-619F23"

BASE COUNT 58444 a 50354 c 49030 g 55779 t 1860 others  
 ORIGIN

Query Match 18.6%; Score 41.8; DB 2; Length 215467;  
 Best Local Similarity 50.2%; Pred. No. 3.9;  
 Matches 103; Conservative 0; Mismatches 102; Indels 0; Gaps 0;

QY 21 CATAAATTTGAATAAATATGAGGCTCCATGATGCTATATGTTTACCTTCAGA 80  
 DB 190461 CGTTTAATTCGTTTATACCTAATATACCTGATACATATGTCAGTTTCTCTCT 190520

QY 81 AGAATATTAGTTTCACTCAGGTTTTTCAAGCTAGGCTGCTCCCAAAAACGAAACAA 140  
 DB 190521 TTTATTTTGGGTTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 190580

QY 141 AACAAAAAACAACTTTTAAAGCTTGTGATGCTGCTACTCATTTGATGCTGCTCTGCTG 200  
 DB 190581 TCCATATTCTCTCTTTTAAATTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTT 190640

QY 201 AATCAATTAGGAATTTTCTTTTTT 225  
 DB 190641 TCCTAGTGTACTCTTTTTTTTTT 190665

Search completed: June 7, 2003, 20:57:10  
 Job time : 1311 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 9, 2003, 08:36:50 ; Search time 1687 Seconds  
(without alignments)  
3353.325 Million cell updates/sec

Title: US-09-765-231A-58  
Perfect score: 225  
Sequence: 1 tgaatggaagtgttccagg.....attagaatttttttttttt 225

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 24791104 seqs, 12571243825 residues

al number of hits satisfying chosen parameters: 49582208

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending Patents NA Main:\*

1: /cgn2\_6/ptodata/1/pna/US06 COMB.seq.\*  
2: /cgn2\_6/ptodata/1/pna/US06 COMB.seq.\*  
3: /cgn2\_6/ptodata/1/pna/US07 COMB.seq.\*  
4: /cgn2\_6/ptodata/1/pna/US08 COMB.seq.\*  
5: /cgn2\_6/ptodata/1/pna/US08 COMB.seq.\*  
6: /cgn2\_6/ptodata/1/pna/US08 COMB.seq.\*  
7: /cgn2\_6/ptodata/1/pna/US08 COMB.seq.\*  
8: /cgn2\_6/ptodata/1/pna/US08 COMB.seq.\*  
9: /cgn2\_6/ptodata/1/pna/US08 COMB.seq.\*  
10: /cgn2\_6/ptodata/1/pna/US08 COMB.seq.\*  
11: /cgn2\_6/ptodata/1/pna/US08 COMB.seq.\*  
12: /cgn2\_6/ptodata/1/pna/US08 COMB.seq.\*  
13: /cgn2\_6/ptodata/1/pna/US08 COMB.seq.\*  
14: /cgn2\_6/ptodata/1/pna/US08 COMB.seq.\*  
15: /cgn2\_6/ptodata/1/pna/US08 COMB.seq.\*  
16: /cgn2\_6/ptodata/1/pna/US08 COMB.seq.\*  
17: /cgn2\_6/ptodata/1/pna/US08 COMB.seq.\*  
18: /cgn2\_6/ptodata/1/pna/US08 COMB.seq.\*  
19: /cgn2\_6/ptodata/1/pna/US08 COMB.seq.\*  
20: /cgn2\_6/ptodata/1/pna/US08 COMB.seq.\*  
21: /cgn2\_6/ptodata/1/pna/US08 COMB.seq.\*  
22: /cgn2\_6/ptodata/1/pna/US08 COMB.seq.\*  
23: /cgn2\_6/ptodata/1/pna/US08 COMB.seq.\*  
24: /cgn2\_6/ptodata/1/pna/US08 COMB.seq.\*  
25: /cgn2\_6/ptodata/1/pna/US08 COMB.seq.\*  
26: /cgn2\_6/ptodata/1/pna/US08 COMB.seq.\*  
27: /cgn2\_6/ptodata/1/pna/US08 COMB.seq.\*  
28: /cgn2\_6/ptodata/1/pna/US08 COMB.seq.\*  
29: /cgn2\_6/ptodata/1/pna/US08 COMB.seq.\*  
30: /cgn2\_6/ptodata/1/pna/US08 COMB.seq.\*  
31: /cgn2\_6/ptodata/1/pna/US08 COMB.seq.\*  
32: /cgn2\_6/ptodata/1/pna/US08 COMB.seq.\*  
33: /cgn2\_6/ptodata/1/pna/US08 COMB.seq.\*  
34: /cgn2\_6/ptodata/1/pna/US08 COMB.seq.\*  
35: /cgn2\_6/ptodata/1/pna/US08 COMB.seq.\*  
36: /cgn2\_6/ptodata/1/pna/US08 COMB.seq.\*  
37: /cgn2\_6/ptodata/1/pna/US08 COMB.seq.\*  
38: /cgn2\_6/ptodata/1/pna/US08 COMB.seq.\*  
39: /cgn2\_6/ptodata/1/pna/US08 COMB.seq.\*  
40: /cgn2\_6/ptodata/1/pna/US08 COMB.seq.\*  
41: /cgn2\_6/ptodata/1/pna/US08 COMB.seq.\*  
42: /cgn2\_6/ptodata/1/pna/US08 COMB.seq.\*  
43: /cgn2\_6/ptodata/1/pna/US08 COMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	225	100.0	225	30	US-09-765-231A-58
2	225	100.0	225	61	US-60-176-523-58
3	225	100.0	320	18	US-09-495-050A-271
4	225	100.0	320	55	US-60-118-318-159
5	214	95.1	1025	61	US-60-172-360-2785
6	196	87.1	1097	25	US-09-649-163-7215
7	196	87.1	1097	25	US-09-652-109-8606
8	183	81.3	860	27	US-09-698-010-152
9	172	76.4	1183	25	US-09-649-163-7215
10	172	76.4	1183	27	US-09-698-010-11426
11	45.2	20.1	1664976	1	PCT-US97-14900A-1
12	45.2	20.1	1664976	13	US-08-916-421-1
13	45.2	20.1	1664976	13	US-08-916-421B-1
14	42.2	18.8	538	81	US-60-377-240-2534
15	41.2	18.3	401	30	US-09-790-483-3984
16	40.8	18.1	397	22	US-09-553-094-9112
17	40	17.8	6809	42	US-10-221-613-173
18	39.6	17.6	1443	22	US-09-577-410-7932
19	39.4	17.5	88098	20	US-09-534-859-279
20	39.4	17.5	88098	31	US-09-803-736-279
21	39.2	17.4	335	31	US-09-821-837-3726

Schnizer, R.  
091765223/page 1  
seq ID 58 w/  
inters

Sequence 16695, A  
Sequence 50276, A  
Sequence 50276, A  
Sequence 50276, A  
Sequence 7217, Ap  
Sequence 429, App  
Sequence 70, Appl  
Sequence 51898, A  
Sequence 22038, A  
Sequence 22038, A  
Sequence 4986, Ap  
Sequence 1, Appli  
Sequence 138837,  
Sequence 138837,  
Sequence 46765, A  
Sequence 58745, A  
Sequence 17932, A  
Sequence 76784, A  
Sequence 292, App  
Sequence 1544, Ap  
Sequence 13175, A  
Sequence 341043,  
Sequence 341043,  
Sequence 8758, Ap

ALIGNMENTS

RESULT 1  
US-09-765-231a-58  
; Sequence 58, Application US/09765231A  
; GENERAL INFORMATION:  
; APPLICANT: Searle/Monsanto  
; APPLICANT: Phippard, Deborah  
; APPLICANT: Vasanthakamur, Geetha  
; APPLICANT: Dotson, Stanton  
; APPLICANT: Ma, Xiao-Jun  
; TITLE OF INVENTION: Osteoarthritis tissue-derived nucleic acids, polypeptides,  
; FILE REFERENCE: SO-3221 PR  
; CURRENT APPLICATION NUMBER: US/09/765, 231A  
; PRIOR FILING DATE: 2001-01-18  
; NUMBER OF SEQ ID NOS: 82  
; SEQ ID NO 58  
; LENGTH: 225  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
09-765-231A-58

Query Match 100.0%; Score 225; DB 30; Length 225;  
Best Local Similarity 100.0%; Pred. No. 1.6e-38;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TGATGGTAAGTTGTTTCAGGCATATAAATTTGAAATAAATATGAGGCTCCATGATGCT 60  
DB 1 TGATGGTAAGTTGTTTCAGGCATATAAATTTGAAATAAATATGAGGCTCCATGATGCT 60  
QY 61 ATATTGGTTTACCTTCAGAGCAATATTAGTTTCACTCAGGTTTTCAGGCTACGCTG 120  
DB 61 ATATTGGTTTACCTTCAGAGCAATATTAGTTTCACTCAGGTTTTCAGGCTACGCTG 120  
QY 121 TCCCCCAAAACGAAACAAACAAACAAACAAACCTTTTAAAGAGTTGATGGCTACTCAT 180  
DB 121 TCCCCCAAAACGAAACAAACAAACAAACAAACCTTTTAAAGAGTTGATGGCTACTCAT 180  
QY 181 TTGATCTGCTCTCTGCTGAATCAATTAGGAATTTTTTTTTTTT 225  
DB 181 TTGATCTGCTCTCTGCTGAATCAATTAGGAATTTTTTTTTTTT 225

Sequence 58, Application US/60176523  
; GENERAL INFORMATION:  
; APPLICANT: Searle/Monsanto  
; APPLICANT: Phippard, Deborah  
; APPLICANT: Vasanthakamur, Geetha  
; APPLICANT: Dotson, Stanton  
; APPLICANT: Ma, Xiao-Jun  
; TITLE OF INVENTION: Osteoarthritis tissue-derived nucleic acids, polypeptides,  
; FILE REFERENCE: SO-3221 PR  
; CURRENT APPLICATION NUMBER: US/60/176, 523  
; PRIOR FILING DATE: 2000-01-18  
; NUMBER OF SEQ ID NOS: 82  
; SEQ ID NO 58  
; LENGTH: 225  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-60-176-523-58

Query Match 100.0%; Score 225; DB 61; Length 225;  
Best Local Similarity 100.0%; Pred. No. 1.6e-38;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TGATGGTAAGTTGTTTCAGGCATATAAATTTGAAATAAATATGAGGCTCCATGATGCT 60  
DB 1 TGATGGTAAGTTGTTTCAGGCATATAAATTTGAAATAAATATGAGGCTCCATGATGCT 60  
QY 61 ATATTGGTTTACCTTCAGAGCAATATTAGTTTCACTCAGGTTTTCAGGCTACGCTG 120  
DB 61 ATATTGGTTTACCTTCAGAGCAATATTAGTTTCACTCAGGTTTTCAGGCTACGCTG 120  
QY 121 TCCCCCAAAACGAAACAAACAAACAAACCTTTTAAAGAGTTGATGGCTACTCAT 180  
DB 121 TCCCCCAAAACGAAACAAACAAACAAACCTTTTAAAGAGTTGATGGCTACTCAT 180  
QY 181 TTGATCTGCTCTCTGCTGAATCAATTAGGAATTTTTTTTTTTT 225  
DB 181 TTGATCTGCTCTCTGCTGAATCAATTAGGAATTTTTTTTTTTT 225

RESULT 3  
US-09-495-050A-271  
; Sequence 271, Application US/09495050A  
; GENERAL INFORMATION:  
; APPLICANT: Roopa, Reddy  
; APPLICANT: Guegler, Karl, J.  
; APPLICANT: Au-Yang, Janice  
; TITLE OF INVENTION: COMPOSITION FOR DETECTION OF GENES ENCODING MEMBRANE-ASSOCIATED P  
; FILE REFERENCE: PA-0013 US  
; CURRENT APPLICATION NUMBER: US/09/495,050A  
; PRIOR FILING DATE: 2000-01-31  
; PRIOR APPLICATION NUMBER: 60/118,318  
; PRIOR FILING DATE: February 1, 1999  
; NUMBER OF SEQ ID NOS: 305  
; SOFTWARE: PERL Program  
; SEQ ID NO 271  
; LENGTH: 320  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No: 4132547CT1  
US-09-495-050A-271

Query Match 100.0%; Score 225; DB 18; Length 320;  
Best Local Similarity 100.0%; Pred. No. 1.7e-38;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TGATGGTAAGTTGTTTCAGGCATATAAATTTGAAATAAATATGAGGCTCCATGATGCT 60  
DB 1 TGATGGTAAGTTGTTTCAGGCATATAAATTTGAAATAAATATGAGGCTCCATGATGCT 60  
QY 61 ATATTGGTTTACCTTCAGAGCAATATTAGTTTCACTCAGGTTTTCAGGCTACGCTG 120

Db 61 ATATTGGTTTACCTTCAGAGGAATATTAGTTTCACTCAGGTTTTTCAAAGCTACGCTG 120  
Qy 121 TCCCCCAAAAGCAAAACAAAAAACAACCTTTTTTAAGAGTTGATGGCTACTCAT 180  
Db 121 TCCCCCAAAAGCAAAACAAAAAACAACCTTTTTTAAGAGTTGATGGCTACTCAT 180  
Qy 181 TTGATCTGCCTCTCTGCTGCTGAATCAATAGGAATTTTTTTTTTTT 225  
Db 181 TTGATCTGCCTCTCTGCTGCTGAATCAATAGGAATTTTTTTTTTTT 225

RESULT 4  
US-60-118-318-159  
; Sequence 159, Application US/60118318  
; GENERAL INFORMATION:  
; APPLICANT: Roopa, Reddy  
; APPLICANT: Guegler, Karl, J.  
; APPLICANT: Au-Young, Janice  
; TITLE OF INVENTION: COMPOSITION FOR DETECTION OF GENES ENCODING MEMBRANE-ASSOCIATED P  
; FILE REFERENCE: PA-0013 P  
; CURRENT APPLICATION NUMBER: US/60/118,318  
; CURRENT FILING DATE: 1999-02-01  
; NUMBER OF SEQ ID NOS: 306  
; SOFTWARE: PERL Program  
; SEQ ID NO 159  
; LENGTH: 320  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: 4132547CT1  
US-60-118-318-159

Query Match 100.0%; Score 225; DB 55; Length 320;  
Best Local Similarity 100.0%; Pred. No. 1.7e-38;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 TGATGGTAAGTTGTTTCAGGCATATAATTTGAAATAAATTTATGAGGCTCCATGATGCT 60  
Db 1 TGATGGTAAGTTGTTTCAGGCATATAATTTGAAATAAATTTATGAGGCTCCATGATGCT 60  
Qy 61 ATATTGGTTTACCTTCAGAGGAATATTAGTTTCACTCAGGTTTTTCAAAGCTACGCTG 120  
Db 61 ATATTGGTTTACCTTCAGAGGAATATTAGTTTCACTCAGGTTTTTCAAAGCTACGCTG 120  
Qy 121 TCCCCCAAAAGCAAAACAAAAAACAACCTTTTTTAAGAGTTGATGGCTACTCAT 180  
Db 121 TCCCCCAAAAGCAAAACAAAAAACAACCTTTTTTAAGAGTTGATGGCTACTCAT 180  
Qy 181 TTGATCTGCCTCTCTGCTGCTGAATCAATAGGAATTTTTTTTTTTT 225  
Db 181 TTGATCTGCCTCTCTGCTGCTGAATCAATAGGAATTTTTTTTTTTT 225

RESULT 5  
US-60-172-360-2785  
; Sequence 2785, Application US/60172360  
; GENERAL INFORMATION:  
; APPLICANT: Morris, MacDonald  
; APPLICANT: Lal, Preeti  
; APPLICANT: Diep, Dinh  
; TITLE OF INVENTION: Method for the Identification of Sequence Polymorphisms Using  
; FILE REFERENCE: Polynucleotide Sequence Databases, and Single Nucleotide Polymor  
; FILE REFERENCE: GX-0007 P  
; CURRENT APPLICATION NUMBER: US/60/172,360  
; CURRENT FILING DATE: 1999-12-16  
; NUMBER OF SEQ ID NOS: 29838  
; SOFTWARE: PERL Program  
; SEQ ID NO 2785  
; LENGTH: 1025  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:

; NAME/KEY: misc feature  
; OTHER INFORMATION: Incey ID No: 115406.1  
US-60-172-360-2785  
Query Match 95.1%; Score 214; DB 61; Length 1025;  
Best Local Similarity 99.6%; Pred. No. 4.6e-36;  
Matches 225; Conservative 0; Mismatches 0; Indels 1; Gaps 1;  
Qy 1 TGATGGTAAGTTGTTTCAGGCATATAATTTGAAATAAATTTATGAGGCTCCATGATGCT 60  
Db 1 TGATGGTAAGTTGTTTCAGGCATATAATTTGAAATAAATTTATGAGGCTCCATGATGCT 60  
Qy 61 ATATTGGTTTACCTTCAGAGGAATATTAGTTTCACTCAGGTTTTTCAAAGCTACGCTG 120  
Db 61 ATATTGGTTTACCTTCAGAGGAATATTAGTTTCACTCAGGTTTTTCAAAGCTACGCTG 120  
Qy 121 TCCCCCAAAAGCAAAACAAAAAACAACCTTTTTTAAGAGTTGATGGCTACTCAT 179  
Db 121 TCCCCCAAAAGCAAAACAAAAAACAACCTTTTTTAAGAGTTGATGGCTACTCA 180  
Qy 180 TTGATCTGCCTCTCTGCTGCTGAATCAATAGGAATTTTTTTTTTTT 225  
Db 181 TTGATCTGCCTCTCTGCTGCTGAATCAATAGGAATTTTTTTTTTTT 226

RESULT 6  
US-09-649-163-7215/c  
; Sequence 7215, Application US/09649163  
; GENERAL INFORMATION:  
; APPLICANT: Holtzman, Douglas A.  
; APPLICANT: Galvin, Katherine A.  
; APPLICANT: Leiby, Kevin R.  
; APPLICANT: Kingsbury, Gillian A.  
; APPLICANT: Welch, Nadine S.  
; APPLICANT: McCarthy, Sean A.  
; APPLICANT: Williamson, Mark  
; APPLICANT: Richardson, Jennifer  
; APPLICANT: MacBeth, Kyle J.  
; APPLICANT: Fraser, Christopher C.  
; APPLICANT: Villevall, Jean-Luc M.G.  
; APPLICANT: Goodearl, Andrew D.J.  
; APPLICANT: Silos-Santiago, Inmaculada  
; APPLICANT: White, David  
; APPLICANT: Pan, Yang  
; APPLICANT: Busfield, Samantha J.  
; APPLICANT: Deeds, James  
; APPLICANT: Lee, John  
; APPLICANT: Shyjan, Andrew W.  
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES  
; FILE REFERENCE: THEREFOR  
; FILE REFERENCE: 1600.1164-001  
; CURRENT APPLICATION NUMBER: US/09/649,163  
; CURRENT FILING DATE: 2000-08-25  
; PRIOR APPLICATION NUMBER: 60/150,608  
; PRIOR FILING DATE: 1999-08-25  
; NUMBER OF SEQ ID NOS: 10535  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7215  
; LENGTH: 1097  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)-(1097)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-649-163-7215

Query Match 87.1%; Score 196; DB 25; Length 1097;  
Best Local Similarity 99.5%; Pred. No. 3.4e-32;  
Matches 207; Conservative 0; Mismatches 0; Indels 1; Gaps 1;  
Qy 19 GCATATAAATTTGAAATAAATTTATGAGGCTCCATGATGCTATATTTGGTTTACCTTCA 78  
|||||

Db 1092 GGCATATAATTTGAAATAAATTAAGGCTCCATGATATGCTATATATGTTTACCTTCA 1033  
Qy 79 GAAGATAATTTAGTTTCACTCAGGTTTTTCAAA-GCTAGCTGTCCCAAAACGAAA 137  
Db 1032 GAAGATAATTTAGTTTCACTCAGGTTTTTCAAAAGGCTAGCTGTCCCAAAACGAAA 973  
Qy 138 CAAAACAAAACAAACCTTTTTTAAGAGTTGATGGCTACTCATTTGATCTGCCTCTCTG 197  
Db 972 CAAAACAAAACAAACCTTTTTTAAGAGTTGATGGCTACTCATTTGATCTGCCTCTCTG 913  
Qy 198 CTGAATCAATAGGAATTTTTTTTTTTT 225  
Db 912 CTGAATCAATAGGAATTTTTTTTTTTT 885

RESULT 7

US-09-652-109-8606/c  
; Sequence 8606, Application US/09652109  
; GENERAL INFORMATION:  
APPLICANT: McCarthy, Sean A.  
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES  
; FILE REFERENCE: 1600.1180-001  
; CURRENT APPLICATION NUMBER: US/09/652,109  
; CURRENT FILING DATE: 2000-08-30  
; PRIOR APPLICATION NUMBER: 60/151,128  
; PRIOR FILING DATE: 1999-08-30  
; NUMBER OF SEQ ID NOS: 10105  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8606  
; LENGTH: 1097  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(1097)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-652-109-8606

Query Match 87.1%; Score 196; DB 25; Length 1097;  
Best Local Similarity 99.5%; Pred. No. 3.4e-32;  
Matches 207; Conservative 0; Mismatches 0; Indels 1; Gaps 1;  
Qy 19 GGCATATAATTTGAAATAAATTAAGGCTCCATGATATGCTATATGTTTACCTTCA 78  
Db 1092 GGCATATAATTTGAAATAAATTAAGGCTCCATGATATGCTATATGTTTACCTTCA 1033  
79 GAAGATAATTTAGTTTCACTCAGGTTTTTCAAA-GCTAGCTGTCCCAAAACGAAA 137  
1032 GAAGATAATTTAGTTTCACTCAGGTTTTTCAAAAGGCTAGCTGTCCCAAAACGAAA 973  
Qy 138 CAAAACAAAACAAACCTTTTTTAAGAGTTGATGGCTACTCATTTGATCTGCCTCTCTG 197  
Db 972 CAAAACAAAACAAACCTTTTTTAAGAGTTGATGGCTACTCATTTGATCTGCCTCTCTG 913  
Qy 198 CTGAATCAATAGGAATTTTTTTTTTTT 225  
Db 912 CTGAATCAATAGGAATTTTTTTTTTTT 885

RESULT 8

US-09-698-010-152  
; Sequence 152, Application US/09698010  
; GENERAL INFORMATION:  
APPLICANT: Williamson, Mark  
APPLICANT: Shyjan, Andrew W.  
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES  
; FILE REFERENCE: 1600.2029-001  
; CURRENT APPLICATION NUMBER: US/09/698,010  
; CURRENT FILING DATE: 2000-10-27  
; PRIOR APPLICATION NUMBER: 60/162,358  
; PRIOR FILING DATE: 1999-10-29

; NUMBER OF SEQ ID NOS: 15684  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 152  
; LENGTH: 860  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(860)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-698-010-152

Query Match 81.3%; Score 183; DB 27; Length 860;  
Best Local Similarity 94.2%; Pred. No. 2e-29;  
Matches 212; Conservative 0; Mismatches 10; Indels 3; Gaps 2;  
Qy 1 TCATGTAAAGTTGTTTTCAGGCATAAAATTTGAAATAAATTAAGGCTCCATGATATGCT 60  
Db 72 TCATGTAAAGTTGTTTTCAGGCATAAAATTTGAAATAAATTAAGGCTCC-CATGATATG 130  
Qy 61 ATATTGGTTTTCACCTTCAGAGAAATATTAGTTTCACTCAGGTTTTTCAAAGCTACGCTG 120  
Db 131 CTTTGGTTTTCACCTTCAGAGAAATATTAGTTTCACTCAGGTTTTTCAAAG--TCGCTG 188  
Qy 121 TCCCCCAAAACGAAACAAACAAACAAACCTTTTAAAGAGTTGATGCTACTCAT 180  
Db 189 TCCCCCAAAACGAAACAAACAAACAAACCTTTTAAAGAGTTGATGCTACTCAT 248  
Qy 181 TTGATCTGCTCTCTGCTGAATCAATTAGGAATTTTTTTTTTTT 225  
Db 249 TTGATCTGCTCTCTGCTGAATCAATTAGGAATTTTTTTTTTTT 293

RESULT 9

US-09-649-163-9422  
; Sequence 9422, Application US/09649163  
; GENERAL INFORMATION:  
APPLICANT: Holtzman, Douglas A.  
APPLICANT: Leiby, Kevin R.  
APPLICANT: Kingsbury, Gillian A.  
APPLICANT: Weich, Nadine S.  
APPLICANT: McCarthy, Sean A.  
APPLICANT: Williamson, Mark  
APPLICANT: Richardson, Jennifer  
APPLICANT: MacBeth, Kyle J.  
APPLICANT: Fraser, Christopher C.  
APPLICANT: Villevall, Jean-Luc M.G.  
APPLICANT: Goodearl, Andrew D.J.  
APPLICANT: Silos-Santiago, Inmaculada  
APPLICANT: White, David  
APPLICANT: Pan Yang  
APPLICANT: Busfield, Samantha J.  
APPLICANT: Deeds, James  
APPLICANT: Lee, John  
APPLICANT: Shyjan, Andrew W.  
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES  
; FILE REFERENCE: 1600.1164-001  
; CURRENT APPLICATION NUMBER: US/09/649,163  
; CURRENT FILING DATE: 2000-08-25  
; PRIOR APPLICATION NUMBER: 60/150,608  
; PRIOR FILING DATE: 1999-08-25  
; NUMBER OF SEQ ID NOS: 10535  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 9422  
; LENGTH: 1183  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-649-163-9422  
Query Match 76.4%; Score 172; DB 25; Length 1183;  
Best Local Similarity 89.7%; Pred. No. 4.9e-27;

Matches 210; Conservative 0; Mismatches 15; Indels 9; Gaps 2;  
QY 1 TGATGGTAAGTTGTTTCAGGCATATAAATTTGAAATAA-----ATTATGAGGCTCCATGAT 55  
Db 75 TGATGGTAAGTTGTTTCAGGCATATAAATTTGAAACATAAATTTGAGGGCTCCAATGAT 134  
QY 56 ATGCTATATTGGTTTTCACCTTCAGA---AGATATTATTAGTTTTCACCTCAGGTTTTCAAA 111  
Db 135 ATGCTATATTGGTTTTCACCTTCAGAAGAAATATTATTAGTTTTCACCTCAGGTTTTCAAA 194  
QY 112 GCTACGCTGTCCTCCCAAAACAAACAAACAAACAAACAAACCTTTTAAAGAGTTGATG 171  
Db 195 GCTACGCTGTCCTCCCAAAACAAACAAACAAACAAACAAACCTTTTAAAGAGTTGATG 254  
QY 172 GCTACTCATTTGATCTGCCTCTCTGCTGTAATCAATTTAGGAATTTTTTTTTTTT 225  
Db 255 GCTACTCATTTGATCTGCCTCTCTGCTGTAATCAATTTAGGAATTTTTTTTTTTT 308

RESULT 10  
US-09-698-010-11426  
Sequence 11426, Application US/09698010  
GENERAL INFORMATION:  
APPLICANT: William, Mark  
APPLICANT: Shvjan, Andrew W.  
TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USBS  
TITLE OF INVENTION: THEREFOR  
FILE REFERENCE: 1600.2029-001  
CURRENT APPLICATION NUMBER: US/09/698,010  
CURRENT FILING DATE: 2000-10-27  
PRIOR APPLICATION NUMBER: 60/162,358  
PRIOR FILING DATE: 1999-10-29  
NUMBER OF SEQ ID NOS: 15684  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 11426  
LENGTH: 1183  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-698-010-11426

Query Match 76.4%; Score 172; DB 27; Length 1183;  
Best Local Similarity 89.7%; Pred. No. 4.9e-27;  
Matches 210; Conservative 0; Mismatches 15; Indels 9; Gaps 2;  
QY 1 TGATGGTAAGTTGTTTCAGGCATATAAATTTGAAATAA-----ATTATGAGGCTCCATGAT 55  
Db 75 TGATGGTAAGTTGTTTCAGGCATATAAATTTGAAACATAAATTTGAGGGCTCCAATGAT 134  
QY 56 ATGCTATATTGGTTTTCACCTTCAGA---AGATATTATTAGTTTTCACCTCAGGTTTTCAAA 111  
Db 135 ATGCTATATTGGTTTTCACCTTCAGAAGAAATATTATTAGTTTTCACCTCAGGTTTTCAAA 194  
QY 112 GCTACGCTGTCCTCCCAAAACAAACAAACAAACAAACCTTTTAAAGAGTTGATG 171  
Db 195 GCTACGCTGTCCTCCCAAAACAAACAAACAAACAAACCTTTTAAAGAGTTGATG 254  
QY 172 GCTACTCATTTGATCTGCCTCTCTGCTGTAATCAATTTAGGAATTTTTTTTTTTT 225  
Db 255 GCTACTCATTTGATCTGCCTCTCTGCTGTAATCAATTTAGGAATTTTTTTTTTTT 308

RESULT 11  
PCT-US97-14900A-1  
Sequence 1, Application PC/TUS9714900A  
GENERAL INFORMATION:  
APPLICANT: THE INSTITUTE FOR GENOMIC RESEARCH  
APPLICANT: 9712 MEDICAL CENTER DRIVE  
APPLICANT: ROCKVILLE, MD 20850  
APPLICANT: UNITED STATES OF AMERICA  
APPLICANT: MICROBIOLOGY DEPARTMENT  
APPLICANT: CHAMPAIGN-URBANA, IL 61801  
APPLICANT: UNITED STATES OF AMERICA  
APPLICANT: SCHOOL OF MEDICINE

APPLICANT: DEPARTMENT OF MOLECULAR BIOLOGY AND GENETICS  
APPLICANT: BALTIMORE, MD 21205  
APPLICANT: UNITED STATES OF AMERICA  
APPLICANT: APPLICANTS/INVENTORS: Bult, Carol J.  
APPLICANT: White, Owen R.  
APPLICANT: Smith, Hamilton O.  
APPLICANT: Woese, Carl R.  
APPLICANT: Venter, J. Craig  
TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic  
TITLE OF INVENTION: Archaeon, Methanococcus jannaschii  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 New York Avenue, NW, Suite 600  
CITY: WASHINGTON  
STATE: DC  
COUNTRY: USA  
ZIP: 20005-3934  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US97/14900A  
FILING DATE: 22-AUG-1997  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/024,428  
FILING DATE: 22-AUG-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Eric K. Steffe  
REGISTRATION NUMBER: 36,688  
REFERENCE/DOCKET NUMBER: 1488.029PC01  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 371-2600  
TELEFAX: (202) 371-2540  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1664976 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
PCT-US97-14900A-1  
Query Match 20.1%; Score 45.2; DB 1; Length 1664976;  
Best Local Similarity 51.0%; Pred. No. 24;  
Matches 107; Conservative 0; Mismatches 103; Indels 0; Gaps 0;  
QY 7 TAAGTTGTTTCAGGCATATAAATTTGAAATATAATTTATGAGGCTCCATGATATGCTATATTG 66  
Db 151403 TAAATCTCTGAGCCCTTCTATTTCATCCATCATTTATTATAGCTACATGTTTGGTAAATG 151462  
QY 67 GTTTTACCTTCAGAGAAATATTAGTTTTCACCTCAGGTTTTCAGAGCTAGGCTGTCCTCCC 136  
Db 151463 TCTTTATCAATAGCCCTTCTTAAATCTTCTCGTAAATTTTAAAACTCCGGAGTTGTCT 151522  
QY 127 AAAAAACGAAACAAACAAACAAACAAACCTTTTAAAGAGTTGATGCTACTCATTTGATC 186  
Db 151523 AAAATCTATAAAATCAATTTACTCTTTTCCAACTCTTTTAATTTGTTTATC 151582  
QY 187 TGCCTCTCTGCTGAATCAATTTAGGAATTT 216  
Db 151583 TTACCCAAAATCCCACTTATTAGGAATTT 151612  
RESULT 12  
US-08-916-421-1  
Sequence 1, Application US/08916421  
GENERAL INFORMATION:  
APPLICANT: Bult, Carol J.  
APPLICANT: White, Owen R.  
APPLICANT: Smith, Hamilton O.

```

; APPLICANT: Woese, Carl R.
; APPLICANT: Venter, J. Craig
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic
; TITLE OF INVENTION: Archaeon, Methanococcus jannaschii
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNER, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 New York Avenue, NW, Suite 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/916,421
; FILING DATE: 22-AUG-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/024,428
; FILING DATE: 22-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Eric K. Steffe
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488.0290001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1664976 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
;
; US-08-916-421-1
;
; Query Match 20.1%; Score 45.2; DB 13; Length 1664976;
; Best Local Similarity 51.0%; Pred. No. 24;
; Matches 107; Conservative 0; Mismatches 103; Indels 0; Gaps 0;
;
; QY 7 TAAGTTGTTTCAGCATATAAATTGAATAAATATGAGGCTCCATGATATGCTATATTG 66
; DB 151403 TAAATCTCGACGCCCTTCTATTTCCATTCATATTATATAGCTACATGTTTGGTAAATTG 151462
;
; 67 GTTTTACCTTCAGAAGAAATTTAGTTTCTACTCAGGTTTTTCAAAGCTACGCTGTCCTCCC 126
; 151463 TCTTTATCAATAGCCTCTCTCTAAATCTCTCGTAAATTTTAAACTCCGGAGTTGTCT 151522
;
; QY 127 AAAAAAGCAACAAACAAACAAACAACTTTTAAAGATTGATGCTACTCATTGATC 186
; DB 151523 AAAAAATATATAAATCAATATTACTCTTTTCCAACTACTCTTAATTTGTTTTTATC 151582
;
; QY 187 TGCTCTCTGCTGAATCAATTAGGAATTT 216
; DB 151583 TTACCCAAATCCCACTTATTAGGAATTT 151612
;
; RESULT 13
; US-08-916-421B-1
; Sequence 1, Application US/08916421B
; GENERAL INFORMATION:
; APPLICANT: Bult et al.
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus
; TITLE OF INVENTION: jannaschii
; FILE REFERENCE: PB275
; CURRENT APPLICATION NUMBER: US/08/916,421B
; CURRENT FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: US 60/024,428
; PRIOR FILING DATE: 1996-08-22
; NUMBER OF SEQ ID NOS: 3
;
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1664976
; TYPE: DNA
; ORGANISM: Methanococcus jannaschii
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (28222)..(28222)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (28257)..(28258)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84773)..(84773)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84808)..(84808)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84812)..(84812)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98120)..(98120)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98159)..(98159)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98239)..(98239)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98266)..(98266)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98343)..(98343)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (103998)..(103998)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (148948)..(148948)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (163385)..(163385)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (191989)..(191989)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (191995)..(191995)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (231980)..(231980)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (234187)..(234187)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (234220)..(234220)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (234814)..(234814)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (309398)..(309398)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (309418)..(309418)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (312837)..(312837)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
```



LOCATION: (312993)..(312993)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (319226)..(319226)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (559167)..(559167)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (559241)..(559241)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (600992)..(600992)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (622708)..(622708)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (657081)..(657081)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (657203)..(657203)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (674435)..(674435)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (682442)..(682442)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (713652)..(713652)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (741684)..(741684)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (779455)..(779455)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (779676)..(779676)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (855539)..(855539)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (871619)..(871619)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1084830)..(1084830)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1096846)..(1096846)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1119881)..(1119881)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1130881)..(1130881)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1310988)..(1310988)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1313224)..(1313224)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1349473)..(1349473)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1349491)..(1349491)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1470091)..(1470091)

OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1569020)..(1569020)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1602912)..(1602912)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1603734)..(1603734)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1637998)..(1637998)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1664854)..(1664854)  
OTHER INFORMATION: n equals a, t, c, or g  
US-08-916-421B-1  
Query Match 20.1%; Score 45.2; DB 13; Length 1664976;  
Best Local Similarity 51.0%; Pred. No. 24;  
Matches 107; Conservative 0; Mismatches 103; Indels 0; Gaps 0;  
QY 7 TAAGTTGTTTCAGGCATAAAATTTGAATAAATATGAGCTCCATGATATGCTATATTG 66  
DB 151403 TAAATCTTCGAGCCCTCTTATTTCCATCCATTATTATAGCTACATGTTTGGTAAATTG 151462  
QY 67 GTTTACCTTCAGAGAATAATTTAGTTTTCACCTCAGGTTTTTCAAAGCTAGCTGTCCCC 126  
DB 151463 TCTTTATCAATAGCCTCTTCTAAATCTTCTCGTAAATTTTAAACTCCGAGTTGCT 151522  
QY 127 AAAAAAGGAAACAAAAACAAAAACCTTTTAAAGTTGATGGCTACTCATTTTGATC 186  
DB 151523 AAAAACTATAAAATCAATTTACTCTTTTCCAAATCTTTTAAATTTGTTTTTATC 151582  
QY 187 TGCTCTCTCTGCTGAATCAATTAGGAATTT 216  
DB 151583 TTACCAAAATCCCACTATTAGGAATTT 151612  
RESULT 14  
US-60-377-240-2534/c  
Sequence 2534, Application US/60377240  
GENERAL INFORMATION:  
APPLICANT: Diggans, James C.  
APPLICANT: Porter, Mark  
APPLICANT: Wei, Tao  
TITLE OF INVENTION: Canine Gene Microarrays for Molecular Toxicology Modeling  
FILE REFERENCE: 44921-5116-PR  
CURRENT APPLICATION NUMBER: US/60/377,240  
CURRENT FILING DATE: 2002-05-03  
PRIOR FILING DATE:  
NUMBER OF SEQ ID NOS: 11109  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2534  
LENGTH: 538  
TYPE: DNA  
ORGANISM: Canis familiaris  
FEATURE:  
OTHER INFORMATION:  
US-60-377-240-2534  
Query Match 18.8%; Score 42.2; DB 81; Length 538;  
Best Local Similarity 55.0%; Pred. No. 33;  
Matches 83; Conservative 0; Mismatches 68; Indels 0; Gaps 0;  
QY 29 TTGAATAAATATGAGCTCCATGATATGCTATATTGTTTACCTTCAGAGAAATATT 88  
DB 537 TTGACATACATTTTGATGTATGTTAGTTAGCTATACAGAAATGTACAAACCAAAGCTGAT 478  
QY 89 TAGTTTCACTCAGGTTTTTCAAAGCTACGCTGTCCCCCAAAAAACAAACAAAAA 148  
DB 477 CAATAATACATAGGTTTTTCAAAAAATTAGGTTTCAATAAAGAAATACAGAGAAAGGAAAT 418

**Qy** 149 AACAACTTTTAAAGATTGATGGTACTCA 179  
||| ||| - ||| | - ||| |||  
**Dδ** 417 AACCACACATCCACAACACTCAGAACTTCTAA 387

```

RESULT 15
US-09-790-482-3984/c
; Sequence 3984, Application US/09790482
; GENERAL INFORMATION:
; APPLICANT: Shvjan, Andrew W.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 1600.2063-001
; CURRENT APPLICATION NUMBER: US/09/790,482
; CURRENT FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: US 60/183,731
; PRIOR FILING DATE: 2000-02-22
; NUMBER OF SEQ ID NOS: 4377
; SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3984
; LENGTH: 401
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(401)
; OTHER INFORMATION: n = A,T,C or G
US-09-790-482-3984

```

[illegible]

Search completed: June 9, 2003, 09:40:39  
 Elapsed time : 1969 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 9, 2003, 08:36:51 ; Search time 2538 Seconds  
(without alignments)  
578.414 Million cell updates/sec

Title: US-09-765-231A-58  
Perfect score: 225  
Sequence: 1 tcatgtagtgggttcagg.....attaggaatttttttttttt 225

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 7578120 seqs, 3262256734 residues

Total number of hits satisfying chosen parameters: 15156240

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: Pending Patents NA New.\*  
1: /cgn2\_6/ptodata/1/pna/PCT\_NEW\_COMB.seq.\*  
2: /cgn2\_6/ptodata/1/pna/US06\_NEW\_COMB.seq.\*  
3: /cgn2\_6/ptodata/1/pna/US07\_NEW\_COMB.seq.\*  
4: /cgn2\_6/ptodata/1/pna/US08\_NEW\_COMB.seq.\*  
5: /cgn2\_6/ptodata/1/pna/US09\_NEW\_COMB.seq.\*  
6: /cgn2\_6/ptodata/1/pna/US09\_NEW\_COMB.seq2.\*  
7: /cgn2\_6/ptodata/1/pna/US09\_NEW\_COMB.seq3.\*  
8: /cgn2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq.\*  
9: /cgn2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq2.\*  
10: /cgn2\_6/ptodata/1/pna/US60\_NEW\_COMB.seq.\*  
11: /cgn2\_6/ptodata/1/pna/US60\_NEW\_COMB.seq2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	225	100.0	320	9	US-10-313-542-271
2	225	100.0	320	9	US-10-313-542A-271
3	45.2	20.1	1664976	6	US-09-692-570-1
4	42.2	18.8	538	1	PCT-US03-13853-2534
5	41.8	18.6	763446	10	US-60-470-166-8943
6	41	18.2	5518	9	US-10-240-453-180
7	39.6	17.6	31312	10	US-60-465-241-52707
8	39.6	17.6	31312	10	US-60-466-412-86559
9	39.2	17.4	60615	8	US-10-427-741-9
10	38.6	17.2	343317	10	US-60-465-241-51757
11	38.6	17.2	343317	10	US-60-466-412-84536
12	38.6	17.2	15105345	5	US-09-948-128-55
13	38	16.9	12393	9	PCT-US02-41414-1544
14	37.8	16.8	96598	1	US-09-837-604A-71071
15	37.6	16.7	522	7	US-09-949-016-12809
16	37.6	16.7	109025	7	US-09-949-016-17567
17	37.6	16.7	109025	7	US-09-947-911-346
18	37.6	16.7	2468502	7	US-09-947-911-26
19	37.6	16.7	2865598	7	US-09-947-911-26
20	37.4	16.6	1524	6	US-09-724-676-31394
21	37.4	16.6	1524	6	US-09-724-676A-31394
22	37.4	16.6	1524	8	US-10-170-235-7647

C 23	37.4	16.6	2586	10	US-60-469-757-297	Sequence 297, App
C 24	37.4	16.6	2586	11	US-60-440-068-297	Sequence 297, App
C 25	37.4	16.6	33519	7	US-09-949-016-17165	Sequence 17165, A
C 26	37.2	16.5	47555	7	US-09-949-016-16549	Sequence 16549, A
C 27	37.2	16.5	49673	7	US-09-949-016-12598	Sequence 12598, A
C 28	37.2	16.5	60583	10	US-60-466-241-53179	Sequence 53179, A
C 29	37.2	16.5	60583	10	US-60-466-412-87261	Sequence 87261, A
C 30	37.2	16.5	218940	7	US-09-949-016-17539	Sequence 17539, A
C 31	37.2	16.5	211522	7	US-09-947-911-63	Sequence 63, Appl
C 32	37	16.4	5506356	5	US-09-948-128-429	Sequence 429, App
C 33	36.8	16.4	601	7	US-09-949-016-205123	Sequence 205123, A
C 34	36.4	16.2	1855	5	US-09-950-083B-12692	Sequence 12692, A
C 35	36.4	16.2	7824	9	US-10-257-166-152	Sequence 152, App
C 36	36.4	16.2	7824	9	US-10-311-455-2112	Sequence 2112, App
C 37	36.4	16.2	29133	10	US-60-466-412-86424	Sequence 86424, A
C 38	36.4	16.2	115823	10	US-60-465-241-51766	Sequence 51766, A
C 39	36.4	16.2	196099	10	US-60-466-412-86709	Sequence 86709, A
C 40	36.4	16.2	495619	7	US-09-947-911-55	Sequence 55, Appl
C 41	36.4	16.2	4925599	5	US-09-948-128-259	Sequence 259, App
C 42	36.2	16.1	5474	9	US-10-311-455-1243	Sequence 1243, Ap
C 43	36.2	16.1	135886	10	US-60-461-762-3375	Sequence 3375, Ap
C 44	36.2	16.1	135886	10	US-60-470-166-8744	Sequence 8744, Ap
C 45	36.2	16.1	176722	5	US-09-948-128-54	Sequence 54, Appl

ALIGNMENTS

RESULT 1  
US-10-313-542-271  
; Sequence 271, Application US/10313542  
; GENERAL INFORMATION:  
; APPLICANT: Roopa, Reddy  
; APPLICANT: Guegler, Karl, J.  
; APPLICANT: Au-Young, Janice  
; TITLE OF INVENTION: COMPOSITION FOR DETECTION OF GENES ENCODING MEMBRANE-ASSOCIATED P  
; FILE REFERENCE: PA-0013 US  
; CURRENT APPLICATION NUMBER: US/10/313,542  
; PRIOR FILING DATE: 2002-12-05  
; PRIOR APPLICATION NUMBER: US/09/495,050  
; PRIOR FILING DATE: 2000-01-31  
; PRIOR APPLICATION NUMBER: 60/118,318  
; PRIOR FILING DATE: 1999-02-01  
; NUMBER OF SEQ ID NOS: 305  
; SOFTWARE: PERL Program  
; SEQ ID NO 271  
; LENGTH: 320  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No: 4132547CT1  
US-10-313-542-271

Query Match	100.0%;	Score 225;	DB 9;	Length 320;
Best Local Similarity	100.0%;	Pred. NO. 2.4e-42;		
Matches	225;	Conservative	0;	Mismatches
Indels	0;	Gaps	0;	
Qy	1	TCATGCTAGTGGTTTCAGGCATAAATTTGAATAAATATGAGGCTCCATGATGCT	60	
Db	1	TCATGCTAGTGGTTTCAGGCATAAATTTGAATAAATATGAGGCTCCATGATGCT	60	
Qy	61	ATATTGGTTTACCTTCAGAGAAATATTAGTTTCACTCAGGTTTTTCAAAAGCTAGCTG	120	
Db	61	ATATTGGTTTACCTTCAGAGAAATATTAGTTTCACTCAGGTTTTTCAAAAGCTAGCTG	120	
Qy	121	TCCCCCAAAAAACGAAAAACAAAAACAACTTTTAAAGAGTTGATGGTACTCAT	180	
Db	121	TCCCCCAAAAAACGAAAAACAAAAACAACTTTTAAAGAGTTGATGGTACTCAT	180	
Qy	181	TTGATCTGCTCTGCTGCTGAATCAATAGGAATTTTTTTTTTTT	225	
Db	181	TTGATCTGCTCTGCTGCTGAATCAATAGGAATTTTTTTTTTTT	225	

```
RESULT 2
US-10-313-542A-271
; Sequence 271, Application US/10313542A
; GENERAL INFORMATION:
; APPLICANT: Roopa, Reddy M.
; APPLICANT: Guegler, Karl, J.
; APPLICANT: Au-Young, Janice K.
; TITLE OF INVENTION: COMPOSITION FOR DETECTION OF GENES ENCODING MEMBRANE-ASSOCIATED
; PROTEINS
; FILE REFERENCE: PA-0013-1 CON
; CURRENT APPLICATION NUMBER: US/10/313,542A
; CURRENT FILING DATE: 2002-12-05
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/118,318
; PRIOR FILING DATE: 1999-02-01
; NUMBER OF SEQ ID NOS: 305
; SOFTWARE: PERL Program
SEQ ID NO 271
; LENGTH: 320
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 4132547CTI
US-10-313-542A-271

; Query Match 100.0%; Score 225; DB 9; Length 320;
; Best Local Similarity 100.0%; Pred. No. 2.4e-42;
; Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGATGGTAAGTTGTTTCAGGCATAAAATTTGAAATAAATTTATGAGGCTCCATGATGCT 60
Db 1 TGATGGTAAGTTGTTTCAGGCATAAAATTTGAAATAAATTTATGAGGCTCCATGATGCT 60

Qy 61 ATATTGGTTTTACTCTCAGAGAAATTTTACTCTCAGGTTTTCAGAGCTACGCTG 120
Db 61 ATATTGGTTTTACTCTCAGAGAAATTTTACTCTCAGGTTTTCAGAGCTACGCTG 120

Qy 121 TCCCCCAAAAAACGAAACAAAAACAAAAACAACTTTTAAAGATTGATGGCTACTCAT 180
Db 121 TCCCCCAAAAAACGAAACAAAAACAAAAACAACTTTTAAAGATTGATGGCTACTCAT 180

Qy 181 TTGATCTGCCTCTCTGCTGAATCAATAGGAATTTTTTTTTTTT 225
Db 181 TTGATCTGCCTCTCTGCTGAATCAATAGGAATTTTTTTTTTTT 225

RESULT 3
US-09-692-570-1
; Sequence 1, Application US/09692570
; GENERAL INFORMATION:
; APPLICANT: Bult et al.
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanoco
; FILE REFERENCE: PB275C1
; CURRENT APPLICATION NUMBER: US/09/692,570
; CURRENT FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: US 60/024,428
; PRIOR FILING DATE: 1996-08-22
; PRIOR APPLICATION NUMBER: US 08/916,421
; PRIOR FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
; LENGTH: 1664976
; TYPE: DNA
; ORGANISM: Methanococcus jannaschii
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (28222)..(28222)
```

```
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (28257)..(28258)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (84773)..(84773)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (84808)..(84808)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (84812)..(84812)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98120)..(98120)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98159)..(98159)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98239)..(98239)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98266)..(98266)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98343)..(98343)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (103998)..(103998)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (148948)..(148948)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (163385)..(163385)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (191989)..(191989)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (191995)..(191995)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (231980)..(231980)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (234187)..(234187)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (234220)..(234220)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (234814)..(234814)
; OTHER INFORMATION: n equals a, t, c, or g
```

FEATURE:  
NAME/KEY: misc feature  
LOCATION: (309398)..(309398)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (309418)..(309418)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (312837)..(312837)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (312993)..(312993)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (319226)..(319226)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (559167)..(559167)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (559241)..(559241)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (600992)..(600992)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (622708)..(622708)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (657081)..(657081)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (657203)..(657203)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (674435)..(674435)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (682442)..(682442)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (713652)..(713652)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (741684)..(741684)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (779455)..(779455)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (779676)..(779676)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (855539)..(855539)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:

NAME/KEY: misc feature  
LOCATION: (871619)..(871619)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1084830)..(1084830)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1096846)..(1096846)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1119881)..(1119881)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1130881)..(1130881)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1310988)..(1310988)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1313224)..(1313224)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1349473)..(1349473)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
Query Match 20.1%; Score 45.2; DB 6; Length 1664976;  
Best Local Similarity 51.0%; Pred.No. 0.35; Indels 0; Gaps 0;  
Matches 107; Conservative 0; Mismatches 103;  
QY 7 TAAGTTGTTTCAGGCATAAAATTTGAAATAAATTTATGAGGCTCCATGATATGCTATATTG 66  
DB 151403 TAAATCTCTGCAGCCCTTCTATTTCATCCATTTATATAGTACATGTTTGGTAAATTG 151462  
QY 67 GTTTTACCTTCAGAGAATATTAGTTTTCACTCAGSTTTTCAAAGCTAGCTGTCCTCC 126  
DB 151463 TCTTTATCAATAGCCCTTCTTAAATCTTCTCGTAAATTTTAAAACTCCGGAGTTGTCT 151522  
QY 127 AAAAAACGAAACAAACAAACAAACAAACCTTTTAAAGAGTTGATGCTACTCATTTGATC 186  
DB 151523 AAAATCTATAAAATCAATTTACTCTTTTCCAACTCTTTAAATTTGTTTTTATC 151582  
QY 187 TGCCTCCTCTGCTGAATCAATTAGGAATTT 216  
DB 151583 TTACCCAAATCCACCTATTAGGAATTT 151612

RESULT 4  
PCT-US03-13853-2534/c  
Sequence 2534, Application PC/TUS0313853  
GENERAL INFORMATION:  
APPLICANT: Diggans, James C.  
APPLICANT: Porter, Mark  
APPLICANT: Wei, Tao  
TITLE OF INVENTION: Canine Gene Microarrays for Molecular Toxicology Modeling  
FILE REFERENCE: 44921-5116-WO  
CURRENT APPLICATION NUMBER: PCT/US03/13853  
CURRENT FILING DATE: 2003-05-05  
PRIOR APPLICATION NUMBER: US 60/377,240  
PRIOR FILING DATE: 2002-05-03  
NUMBER OF SEQ ID NOS: 11109  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2534  
LENGTH: 538  
TYPE: DNA

```
; ORGANISM: Canis familiaris
; FEATURE:
; OTHER INFORMATION:
PCT-US03-13853-2534

Query Match      18.8%; Score 42.2; DB 1; Length 538;
Best Local Similarity 55.0%; Pred. No. 2;
Matches 83; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 29 TTGAATAAATATGAGGCTCCATGATATGCTATATGTTTACCTTCAGAGATATT 88
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 537 TTGACATACATTTTGATGTATGATGTTAGGCTATACAGATGTACAAACCAAGCTGTAT 478
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 89 TAGTTTCACCTCAGGTTTTTCAAAAGCTAGCGTGTCCCCCAAAAAACGAAACAAACAAAA 148
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 477 CATATACATAGGTTTTTCAAAAATAGGTTTCAATNAAGATACAGAGAAAGGAAT 418
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 149 AACAACTTTTTTAAGATGATGGCTACTCA 179
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
417 AACCAACATCCAACTCAGAACTTCTAA 387

RESULT 5
US-60-470-166-8943
; Sequence 8943, Application US/60470166
; GENERAL INFORMATION:
; APPLICANT: CRUPE, Andrew
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001464
; CURRENT APPLICATION NUMBER: US/60/470,166
; CURRENT FILING DATE: 2003-05-09
; NUMBER OF SEQ ID NOS: 68617
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 8943
; LENGTH: 763446
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(763446)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-60-470-166-8943

Query Match      18.6%; Score 41.8; DB 10; Length 763446;
Best Local Similarity 50.2%; Pred. No. 2.2;
Matches 103; Conservative 0; Mismatches 102; Indels 0; Gaps 0;

21 CATAAAATTTGAAATAAATATGAGGCTCCATGATGCTATATGTTTACCTTCAGA 80
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 102622 CGTTTAATTCGTTTATTAATTAATCTACTCATCATATGATGAGTTTTCCTGT 102681
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 81 AGAATATTTAGTTTCTACTAGGTTTTTCAAGCTAGCGTGTCCCAAAAAACGAAACAA 140
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 102682 TTTATTTGGGTTTCTTTGCTCTCTTCCCTTCTCTATATTAAGGAATGT 102741
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 141 AACAAAAACACACCTTTTAAAGTTGATGGCTACTCATTTGATGCTCTCTCTGCTG 200
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 102742 TCCATATTTCTCTTTTAAAAATTTTGGTTGTTCAAGCTGATGATCATATTTCTGTC 102801
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 201 AATCAATTAGGAATTTTTTTTTTTT 225
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 102802 TCCTAGTGATTAATCTTTTTTTTTT 102826
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 6
US-10-240-453-180/c
; Sequence 180, Application US/10240453
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
```

```
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA
; TITLE OF INVENTION: Transcription
; TITLE OF INVENTION: By Means of Assessing the Methylation Status of Genes Associated
; TITLE OF INVENTION: With DNA Transcription
; FILE REFERENCE: 5013.1009
; CURRENT APPLICATION NUMBER: US/10/240,453
; CURRENT FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: PCT/EP01/03973
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 350
; SEQ ID NO 180
; LENGTH: 5518
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-240-453-180

Query Match      18.2%; Score 41; DB 9; Length 5518;
Best Local Similarity 51.4%; Pred. No. 3.6;
Matches 95; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY 7 TAAGTTGTTTCAGGCATAAATTTGAAATAAATATGAGGCTCCATGATATGCTATATTG 66
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 4088 TATATTTTTTAAACACAAAATTTTCATTTAACTATAAACTCCTCTAAAATATTAAACGAA 4029
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 67 GTTTTACCTTCAGAAGATATTTAGTTTCACTCAGGTTTTTCAAGCTACGCTGTCCTCC 126
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 4028 TTACTTAATTAATACTAAATAATAATACTTCAAACTCTTAAAAACTAATTACTCTAA 3969
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 127 AAAAAACGAAACAAACAAAAACAACTTTTAAAGAGTTGATGCTACTCATTTGATC 186
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3968 AATAAAAATACTTAATAATAATAATACTATTTTAAATAAATACTTAATCTATT 3909
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 187 TGCCT 191
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3908 TAACT 3904
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 7
US-60-465-241-52707
; Sequence 52707, Application US/60465241
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: BEGOVICH, Ann
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001468
; CURRENT APPLICATION NUMBER: US/60/465,241
; CURRENT FILING DATE: 2003-04-23
; NUMBER OF SEQ ID NOS: 258418
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 52707
; LENGTH: 31312
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(31312)
; OTHER INFORMATION: n = A,T,C or G
US-60-465-241-52707

Query Match      17.6%; Score 39.6; DB 10; Length 31312;
Best Local Similarity 58.5%; Pred. No. 7.3;
Matches 69; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
```

QY 108 CAAAGCTACGCTGCTCCCAAAACGAAACAAACAAACAAACAAACAAACCTTTTAAAGAGTT 167  
Db 29922 CAGAGCAAGACTGTCTCAAAAAAAGAAAAAAGAAAAAATTCAAATTTTTCGCTT 29981  
QY 168 GATGGCTACTCATTTGATCTGCTCTGCTGAATCAATTAGGAATTTTTTTTTT 225  
Db 29982 GACCACTGACTATTCAGTTTCTCTATCTTGTGTGAGCAAGTTGTGGTAGTGTCT 30039

RESULT 8  
US-60-466-412-86559  
; Sequence 86559, Application US/60466412  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele  
; APPLICANT: IAKOUBOVA, Olga  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001466  
; CURRENT APPLICATION NUMBER: US/60/466,412  
; CURRENT FILING DATE: 2003-04-30  
; NUMBER OF SEQ ID NOS: 429241  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 86559  
; LENGTH: 31312  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)...(31312)  
; OTHER INFORMATION: n = A,T,C or G  
US-60-466-412-86559

Query Match 17.6%; Score 39.6; DB 10; Length 31312;  
Best Local Similarity 58.5%; Pred. No. 7.3;  
Matches 69; Conservative 0; Mismatches 49; Indels 0; Gaps 0;  
QY 108 CAAAGCTACGCTGCTCCCAAAACGAAACAAACAAACAAACAAACCTTTTAAAGAGTT 167  
Db 29922 CAGAGCAAGACTGTCTCAAAAAAAGAAAAAAGAAAAAATTCAAATTTTTCGCTT 29981  
QY 168 GATGGCTACTCATTTGATCTGCTCTGCTGAATCAATTAGGAATTTTTTTTTT 225  
Db 29982 GACCACTGACTATTCAGTTTCTCTATCTTGTGTGAGCAAGTTGTGGTAGTGTCT 30039

RESULT 9  
US-10-427-741-9  
; Sequence 9, Application US/10427741  
; GENERAL INFORMATION:  
; APPLICANT: Tohyama, Masaya  
; APPLICANT: Yamashita, Toshihide  
; TITLE OF INVENTION: COMPOSITION AND METHOD FOR NERVE REGENERATION  
; FILE REFERENCE: 59150-8023  
; CURRENT APPLICATION NUMBER: US/10/427,741  
; CURRENT FILING DATE: 2003-04-30  
; PRIOR FILING DATE: 2003-03-28  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 9  
; LENGTH: 60615  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-10-427-741-9

Query Match 17.4%; Score 39.2; DB 8; Length 60615;  
Best Local Similarity 62.0%; Pred. No. 8.9;  
Matches 62; Conservative 0; Mismatches 38; Indels 0; Gaps 0;  
QY 74 CTTCAAGATATTTAGTTTCAAGCTACGCTGTTCCCAACAAAC 133  
Db 759 CTACAGAGTGAGTCCAGGACCCAGGGGTATACAGAGAAACCTGTCTGTGTAACAC 818

QY 134 GAAACAAACAAACAAACCTTTTAAAGAGTTGATGCG 173  
Db 819 AAAACAAACAAACAAACAAACAAACAAACAAACATGCG 858

RESULT 10  
US-60-465-241-51757  
; Sequence 51757, Application US/60465241  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele  
; APPLICANT: BEGOVICH, Ann  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001468  
; CURRENT APPLICATION NUMBER: US/60/465,241  
; CURRENT FILING DATE: 2003-04-23  
; NUMBER OF SEQ ID NOS: 258418  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 51757  
; LENGTH: 343317  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)...(343317)  
; OTHER INFORMATION: n = A,T,C or G  
US-60-465-241-51757

Query Match 17.2%; Score 38.6; DB 10; Length 343317;  
Best Local Similarity 61.1%; Pred. No. 12;  
Matches 80; Conservative 0; Mismatches 49; Indels 2; Gaps 1;  
QY 7 TAAGTTGTTTCAGGCATAAATTTGAAATAAATATGAGCTCCATGATGCTATATTG 66  
Db 185157 TTATTTCTTTGAGGTAAACAATTTCAAGAAATAATGATTAATTAATTAATA 185216  
QY 67 GTTTTACCTTCAGAGAATAATTTAGT--TTCACTCAGGTTTTTCAAGCTACGCTGCTCC 124  
Db 185217 ATCTGAAGACAAATGAATCTTTAGGAATCTACTCATTATTTTAAAGATATGCTATCAG 185276  
QY 125 CCAAAAACGA 135  
Db 185277 CCATATATCAA 185287

RESULT 11  
US-60-466-412-84536  
; Sequence 84536, Application US/60466412  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele  
; APPLICANT: IAKOUBOVA, Olga  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001466  
; CURRENT APPLICATION NUMBER: US/60/466,412  
; CURRENT FILING DATE: 2003-04-30  
; NUMBER OF SEQ ID NOS: 429241  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 84536  
; LENGTH: 343317  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)...(343317)  
; OTHER INFORMATION: n = A,T,C or G  
US-60-466-412-84536

Query Match 17.2%; Score 38.6; DB 10; Length 343317;  
Best Local Similarity 61.1%; Pred. No. 12;  
Matches 80; Conservative 0; Mismatches 49; Indels 2; Gaps 1;





```

; APPLICANT: Bougri, Olegs
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; FILE OF INVENTION: Plants
; FILE REFERENCE: 38-21(51892)B
; CURRENT APPLICATION NUMBER: US/09/837,604A
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: US 60/197,872
; PRIOR FILING DATE: 2000-04-19
; NUMBER OF SEQ ID NOS: 81288
; SEQ ID NO 71071
; LENGTH: 522
; TYPE: DNA
; ORGANISM: Oryza sativa nipponbare
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(522)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3598-060-Q1-N1-F10
; US-09-837-604A-71071

```

```

Query Match      16.7%; Score 37.6; DB 7; Length 522;
Best Local Similarity 55.3%; Pred. No. 23;
Matches 73; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 22 ATAAATTTGAAATAAATTATGAGGCTCCATGATATGCTATATTTGGTTTACCTTCAGAA 81
Db 150 AAAAAATATTTTATTAATTTATTTTTCCTCCCTTTGGGGTTTTTTTTTTTCTTG 91

QY 82 GAATATTTAGTTTCACTCAGCTTTTCAAGCTACGCTGTCCTCCCAAAACGAAACAAA 141
Db 90 TAAATTTTTTTTAAACCGTTTTTAAATGTTAGGGCCCCCCCCCAAAAAAAAAA 31

QY 142 ACAAAAAACAA 153
Db 30 AAAAAAAAAA 19

```

Search completed: June 9, 2003, 11:23:37  
Job time : 3591 secs

**THIS PAGE BLANK (USPTO)**